



**TUGAS AKHIR – SS141501**

***MODEL OUTPUT STATISTICS DENGAN CART  
DAN RANDOM FORESTS UNTUK PRAKIRAAN  
CURAH HUJAN HARIAN***

**NURHAYATI  
NRP 1315 105 023**

**Dosen Pembimbing  
Dr. Sutikno, M.Si**

**PROGRAM STUDI SARJANA  
DEPARTEMEN STATISTIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA 2017**





**FINAL PROJECT – SS141501**

***MODEL OUTPUT STATISTICS BY CART  
AND RANDOM FORESTS FOR DAILY  
PRECIPITATION PREDICTION***

**NURHAYATI  
NRP 1315 105 023**

**Dosen Pembimbing  
Dr. Sutikno, M.Si**

**UNDERGRADUATE PROGRAMME  
DEPARTMENT OF STATISTICS  
FACULTY OF MATHEMATICS AND NATURAL SCIENCES  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA 2017**



## LEMBAR PENGESAHAN

### MODEL OUTPUT STATISTICS DENGAN CART DAN RANDOM FORESTS UNTUK PRAKIRAAN CURAH HUJAN HARIAN

#### TUGAS AKHIR

Diajukan Untuk Memenuhi Salah Satu Syarat  
Memperoleh Gelar Sarjana Sains  
pada

Program Studi Sarjana Departemen Statistika  
Fakultas Matematika dan Ilmu Pengetahuan Alam  
Institut Teknologi Sepuluh Nopember

Oleh :

**Nurhayati**

NRP. 1315 105 023

Disetujui oleh Pembimbing:

Dr. Sutikno, M.Si

NIP. 19710313 199702 1 001

(  )



SURABAYA, JULI 2017



# **MODEL OUTPUT STATISTICS DENGAN CART DAN RANDOM FORESTS UNTUK PRAKIRAAN CURAH HUJAN HARIAN**

**Nama Mahasiswa : Nurhayati**  
**NRP : 1315 105 023**  
**Departemen : Statistika FMIPA-ITS**  
**Dosen Pembimbing : Dr. Sutikno, M.Si**

## **Abstrak**

*Cuaca memiliki pengaruh yang sangat penting terhadap kehidupan manusia. Upaya meminimalkan dampak bencana akibat cuaca/iklim maka dibutuhkan informasi prakiraan cuaca/iklim. Khususnya pada bidang transportasi membutuhkan informasi cuaca pendek yang sangat cepat dan tepat. Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) sebagai lembaga resmi pemerintah yang bertugas memberikan informasi cuaca. Salah satu informasi cuaca jangka pendek yang cukup penting adalah curah hujan. BMKG telah berupaya melakukan pemodelan cuaca jangka pendek dengan memanfaatkan data komponen cuaca Numerical Weather Prediction (NWP). Salah satu upaya untuk mengoptimalkan pemanfaatan output model NWP adalah dengan Model Output Statistics (MOS), karena prakiraan NWP masih bias. Metode yang akan digunakan untuk pendekatan MOS adalah Classification and Regression Tree (CART) dan random forests.*

*Tujuan penelitian ini adalah mengetahui hasil ketepatan klasifikasi curah hujan harian melalui metode tersebut. Hasil penelitian menunjukkan bahwa ketepatan klasifikasi untuk metode CART pada Stasiun Meteorologi Juanda, Stasiun Meteorologi Ngurai Rai dan pada Stasiun Meteorologi Soekarno-Hatta (Soetta) yaitu: 46.67%, 73.33% dan 66.67%. Sedangkan untuk metode random forests yaitu: 53.33%, 53.33% dan 70.00%. Metode random forests mampu meningkatkan akurasi hasil prakiraan curah hujan harian.*

**Kata Kunci: CART, Curah Hujan Harian, MOS, NWP, Random Forests**

*(halaman ini sengaja dikosongkan)*



# ***MODEL OUTPUT STATISTICS BY CART AND RANDOM FORESTS FOR DAILY PRECIPITATION PREDICTION***

<b>Student Name</b>	<b>: Nurhayati</b>
<b>NRP</b>	<b>: 1315 105 023</b>
<b>Department</b>	<b>: Statistics FMIPA ITS</b>
<b>Supervisor</b>	<b>: Dr. Sutikno, M.Si</b>

## **Abstract**

*Weather has a very important influence on human life. The efforts to minimize the impact of disasters which caused by weather / climate therefor weather / climate forecasting information is required. Especially in the transportation field which requires a fast and precise of short weather information. Indonesian agency for Meteorological, Climatological and Geophysics (BMKG) as the official government agency is in charge of providing weather information. One of the important short-term weather information is precipitation. BMKG has attempted short-term weather modeling using Numerical Weather Prediction (NWP) weather component data. One of efforts to optimize the utilization of NWP model output is by Model Output Statistics (MOS), as NWP forecasts are still biased. The methods to be used for the MOS approach are Classification and Regression Tree (CART) and random forests.*

*The purpose of this research is to know the accuracy result of daily precipitation classification through the method. The results showed that the accuracy of classification for CART method at Juanda Meteorological Station, Ngurai Rai Meteorology Station and Soekarno-Hatta Meteorological Station (Soetta) are 46.67%, 73.33% and 66.67%. Whereas for random forests method are: 53.33%, 53.33% and 70.00%. Random forests method can improve the accuracy of daily precipitation forecast results.*

**Keywords :** *CART, precipitation daily, MOS, NWP, Random Forest.*

*(halaman ini sengaja dikosongkan)*

## KATA PENGANTAR

Puji syukur penulis panjatkan atas kehadiran Allah SWT yang telah memberikan rahmat, hidayah, nikmat, dan petunjuk-Nya serta tak lupa Sholawat dan salam kepada Rasulullah SAW sehingga penulisan Tugas Akhir dengan judul

***“Model Output Statistics dengan CART dan Random Forests untuk Prakiraan Curah Hujan Harian”***

dapat terselesaikan dengan baik dan tepat pada waktunya. Selama proses penyusunan hingga penyelesaian Tugas Akhir ini tidak lepas dari doa, saran, bimbingan serta dukungan dari berbagai pihak. Oleh karena itu, penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada :

1. Bapak Dr. Sutikno, M.Si selaku dosen pembimbing yang telah bersedia meluangkan waktu untuk memberi bimbingan, ilmu, nasihat beserta saran kepada penulis demi kesempurnaan Tugas Akhir ini.
2. Bapak Dr. Suhartono selaku Ketua Departemen Statistika FMIPA ITS yang telah menyediakan fasilitas guna kelancaran pengerjaan Tugas Akhir ini.
3. Ibu Dr. Santi Wulan Purnami, S.Si, M.Si selaku dosen penguji atas segala kritikan dan saran yang sangat membangun.
4. Bapak Dr. rer.pol. Heri Kuswanto selaku dosen penguji dan dosen wali yang telah memberikan motivasi, saran dan koreksi dalam pengerjaan Tugas Akhir ini.
5. Seluruh dosen Departemen Statistika yang telah memberikan ilmu selama penulis menempuh pendidikan Lintas Jalur.
6. Seluruh staf Tata Usaha dan karyawan Departemen Statistika yang telah membantu kelancaran dalam pelaksanaan kegiatan perkuliahan.
7. Bapak H. Munir dan Ibu Hj. Syarifah yang selalu mendoakan, memberikan dukungan, saran, dan semua kasih sayang yang diberikan, semua sepupu dan keluarga

besar yang telah memberi semangat terus menerus. Semoga penulis bisa menjadi seseorang yang selalu membanggakan kalian.

8. Teman-teman seperjuangan Binti Fatmawati, Silviana Febryani, Rosdiana Sirait dan Widya Umroatun S yang telah memberikan bantuan dan motivasi selama mengerjakan Tugas Akhir ini.
9. Teman-teman satu tim dan satu bimbingan dalam mengerjakan Tugas Akhir Vira Oktavia, Fachrunisah, Partini Ningsih, Nazmatus Sahiroh dan Vella yang telah berbagi ilmu dan cerita.
10. Teman-teman LJ 2015 dan rekan-rekan seperjuangan Lab Lingkungan dan Kesehatan yang tidak dapat disebutkan satu persatu.
11. Teman-teman GW 19 yang telah memberikan dukungan semangat dan doa untuk kelancaran Tugas Akhir ini, serta pengalaman yang berharga bagi kehidupan penulis.
12. Keluarga  $\sum 23$  yang telah memberikan banyak kenangan dari awal hingga saat ini.
13. Dan terakhir kepada semua pihak yang belum bisa disebutkan dan berkontribusi lebih, baik melalui dukungan moril maupun bantuan secara nyata demi suksesnya Tugas Akhir ini.

Penulis menyadari bahwa Tugas Akhir ini masih jauh dari sempurna. Oleh sebab itu, diharapkan saran dan kritik yang membangun dari berbagai pihak untuk dijadikan bahan pertimbangan dalam penyusunan laporan berikutnya. Semoga dapat memberi manfaat kepada penulis dan juga pembaca.

Surabaya, Juli 2017

Penulis

# DAFTAR ISI

	Halaman
<b>HALAMAN JUDUL</b> .....	i
<b>LEMBAR PENGESAHAN</b> .....	iii
<b>Abstrak</b> .....	v
<b>Abstract</b> .....	vii
<b>KATA PENGANTAR</b> .....	ix
<b>DAFTAR ISI</b> .....	xi
<b>DAFTAR GAMBAR</b> .....	xiii
<b>DAFTAR TABEL</b> .....	xv
<b>DAFTAR LAMPIRAN</b> .....	xvii
<b>BAB 1 PENDAHULUAN</b>	
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	3
1.3 Tujuan.....	4
1.4 Manfaat.....	4
1.5 Batasan Masalah.....	4
<b>BAB 2 TINJAUAN PUSTAKA</b>	
2.1 <i>Principal Component Analysis (PCA)</i> .....	5
2.2 <i>Classification and Regression Trees</i> .....	6
2.3 <i>Random Forests</i> .....	11
2.3 Ukuran Ketepatan Klasifikasi .....	13
2.4 <i>Model Output Statistics</i> .....	14
2.5 <i>Numerical Weather Prediction</i> .....	17
2.6 Curah Hujan .....	18
<b>BAB III METODOLOGI PENELITIAN</b>	
3.1 Sumber Data .....	19
3.2 Variabel Penelitian .....	19
3.3 Struktur Data .....	22
3.4 Langkah Analisis .....	23
<b>BAB IV ANALISIS DAN PEMBAHASAN</b>	
4.1 Deskripsi Curah Hujan .....	27
4.2 Reduksi Dimensi Data Numerical Weather Prediction	28
4.3 Penyusunan Model dengan CART .....	30

4.3.1	Klasifikasi Curah Hujan Stamet Juanda .....	30
4.3.2	Klasifikasi Curah Hujan Stamet Ngurah Rai .....	37
4.3.3	Klasifikasi Curah Hujan Stamet Soetta .....	42
4.4	Penyusunan Model <i>Random Forests</i> .....	48
4.5	Kinerja Model.....	50
<b>BAB V KESIMPULAN DAN SARAN</b>		
5.1	Kesimpulan.....	53
5.2	Saran .....	53
<b>DAFTAR PUSTAKA</b>		
<b>LAMPIRAN</b>		

## DAFTAR GAMBAR

	Halaman
<b>Gambar 2.1</b> Ilustrasi Pohon Klasifikasi .....	6
<b>Gambar 2.2</b> Algoritma <i>Random Forests</i> .....	13
<b>Gambar 3.1</b> Pengukuran NWP dalam Grid 3x3.....	22
<b>Gambar 3.2</b> <i>Flowchart</i> Langkah Analisis .....	26
<b>Gambar 4.1</b> Plot Relative Cost Klasifikasi Curah Hujan Stamet Juanda .....	32
<b>Gambar 4.2</b> Split Plot Pohon Optimal Stamet Juanda .....	35
<b>Gambar 4.3</b> Plot Relative Cost Klasifikasi Curah Hujan Stamet Ngurai Rai.....	38
<b>Gambar 4.4</b> Plot Relative Cost Klasifikasi Curah Hujan Stamet Soetta .....	43
<b>Gambar 4.5</b> Split Plot Pohon Optimal Stamet Soetta .....	45

*(halaman ini sengaja dikosongkan)*



## DAFTAR TABEL

	Halaman
<b>Tabel 2.1</b> <i>Crosstab</i> Ketepatan Klasifikasi .....	14
<b>Tabel 2.2</b> Klasifikasi Intensitas Curah Hujan .....	18
<b>Tabel 3.1</b> Wilayah Stasiun Pengamatan .....	19
<b>Tabel 3.2</b> Klasifikasi Curah Hujan Menurut Intensitasnya.....	20
<b>Tabel 3.3</b> Variabel NWP WRF-EMS .....	20
<b>Tabel 3.4</b> Struktur Data Penelitian.....	23
<b>Tabel 4.1</b> Persentase Kejadian Hujan .....	27
<b>Tabel 4.2</b> <i>Eigenvalue</i> dan Kumulatif Keragaman Variabel hujan tot Stamet Juanda .....	28
<b>Tabel 4.3</b> <i>Eigenvalue</i> dan Keragaman PC Stamet Juanda .....	28
<b>Tabel 4.4</b> <i>Eigenvalue</i> dan Keragaman PC Stamet Ngurai Rai....	29
<b>Tabel 4.5</b> <i>Eigenvalue</i> dan Keragaman PC Stamet Soetta .....	30
<b>Tabel 4.6</b> Skor Variabel Prediktor dalam Pembentukan Pohon Klasifikasi Maksimal Stamet Juanda .....	31
<b>Tabel 4.7</b> Pembentukan Pohon Klasifikasi Stamet Juanda.....	32
<b>Tabel 4.8</b> Skor Variabel Prediktor dalam Pembentukan Pohon Klasifikasi Optimal Stamet Juanda .....	33
<b>Tabel 4.9</b> Kelas Curah Hujan Stamet Juanda pada Masing- masing Simpul Terminal .....	34
<b>Tabel 4.10</b> Ketepatan Klasifikasi Data <i>Learning</i> Stamet Juanda	36
<b>Tabel 4.11</b> Ketepatan Klasifikasi Data <i>Testing</i> Stamet Juanda ..	37
<b>Tabel 4.12</b> Skor Variabel Prediktor dalam Pembentukan Pohon Klasifikasi Maksimal Stamet Ngurai Rai .....	37
<b>Tabel 4.13</b> Pembentukan Pohon Klasifikasi Stamet Ngurai Rai	39
<b>Tabel 4.14</b> Skor Variabel Prediktor dalam Pembentukan Pohon Klasifikasi Optimal Stamet Ngurai Rai .....	40
<b>Tabel 4.15</b> Kelas Curah Hujan Stamet Ngurai Rai pada Masing- masing Simpul Terminal .....	40
<b>Tabel 4.16</b> Ketepatan Klasifikasi Data <i>Learning</i> Ngurai Rai ....	41
<b>Tabel 4.17</b> Ketepatan Klasifikasi Data <i>Testing</i> Ngurai Rai .....	42

<b>Tabel 4.18</b> Skor Variabel Prediktor dalam Pembentukan Pohon Klasifikasi Maksimal Stamet Soetta.....	42
<b>Tabel 4.19</b> Pembentukan Pohon Klasifikasi Stamet Soetta .....	44
<b>Tabel 4.20</b> Skor Variabel Prediktor dalam Pembentukan Pohon Klasifikasi Optimal Stamet Soetta.....	44
<b>Tabel 4.21</b> Kelas Curah Hujan Stamet Soetta pada Masing-masing Simpul Terminal .....	46
<b>Tabel 4.22</b> Ketepatan Klasifikasi Data <i>Learning</i> Stamet Soetta.	47
<b>Tabel 4.23</b> Ketepatan Klasifikasi Data <i>Testing</i> Stamet Soetta....	48
<b>Tabel 4.24</b> Hasil Klasifikasi <i>Random Forests</i> Stamet Juanda ....	48
<b>Tabel 4.25</b> Hasil Klasifikasi <i>Random Forests</i> Stamet NRai .....	49
<b>Tabel 4.26</b> Hasil Klasifikasi <i>Random Forests</i> Stamet Soetta .....	49
<b>Tabel 4.27</b> Perbandingan Hasil Klasifikasi antara CART dan <i>Random Forests</i> .....	50

## DAFTAR LAMPIRAN

	Halaman
<b>Lampiran 1.</b> Data Observasi Curah Hujan.....	57
<b>Lampiran 2.</b> Data NWP Stamet Juanda .....	58
<b>Lampiran 3.</b> Data NWP Stamet Ngurai Rai.....	64
<b>Lampiran 4.</b> Data NWP Stamet Soetta .....	70
<b>Lampiran 5.</b> Output Reduksi PCA Stamet Juanda.....	76
<b>Lampiran 6.</b> Output Reduksi PCA Stamet Ngurai Rai .....	84
<b>Lampiran 7.</b> Output Reduksi PCA Stamet Soetta.....	84
<b>Lampiran 8.</b> Output CART Stamet Juanda.....	100
<b>Lampiran 9.</b> Output CART Stamet Ngurai Rai .....	102
<b>Lampiran 10.</b> Output CART Stamet Soetta.....	105
<b>Lampiran 11.</b> Program R Untuk Random Forests .....	107

*(halaman ini sengaja dikosongkan)*

# **BAB 1**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Saat ini pengguna informasi cuaca jangka pendek menuntut untuk memperoleh informasi cuaca secara cepat dan akurat. Informasi cuaca pendek adalah prakiraan cuaca untuk periode 1-7 hari kedepan. Bidang transportasi membutuhkan informasi kondisi cuaca sebelum armada tersebut akan berangkat ke tempat tujuan. Seiring dengan pesatnya penerbangan antar pulau dan antar negara, membutuhkan informasi kondisi cuaca lebih cepat, akurat, tepat waktu dan sekaligus mampu memberikan peringatan dini untuk keselamatan transportasi.

Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) sebagai lembaga resmi pemerintah yang bertugas memberikan informasi cuaca. Salah satu informasi cuaca jangka pendek yang cukup penting adalah curah hujan. Mulai tahun 2004, BMKG telah melakukan penelitian untuk prakiraan cuaca jangka pendek dengan memanfaatkan data komponen cuaca *Numerical Weather Prediction* (NWP). Namun NWP menghasilkan ramalan yang bias terutama untuk lokasi dengan topografi berbukit dan vegetasi yang kompleks. Hal ini dikarenakan NWP memberikan informasi cuaca yang bersifat homogen dan diukur pada skala global. Akibatnya, jika model NWP digunakan untuk meramal cuaca lokasi (dimensi rendah) akan menghasilkan ramalan cuaca yang bias (Wilks, 2006). Oleh karena itu, perlu dilakukan pemrosesan secara statistik (*statistical post processing*) untuk mengoptimalkan pemanfaatan output NWP. Salah satu upaya untuk mengoptimalkan pemanfaatan output model NWP adalah dengan *Model Output Statistics* (MOS).

MOS merupakan model yang menghubungkan antara variabel respon (hasil observasi cuaca) dan variabel prediktor (parameter NWP) dengan model berbasis regresi (Nichols, 2008). MOS menentukan hubungan statistik antara variabel prediktor

dan variabel respon NWP pada beberapa proyeksi waktu (Glahn & Lowry, 1972). Metode MOS dengan berbagai pendekatan regresi telah banyak digunakan untuk prakiraan cuaca jangka pendek, diantaranya: Multivariate Adaptive Regression Spline: Priambudi (2006); regresi berganda: Wardani (2010); *Regresi Partial Least Square*: Sari (2013); *multivariate regression*: Kusumawardani (2014).

Data NWP diambil dengan 9 grid pengukuran di setiap lokasi untuk masing-masing variabel. Dalam pengukuran variabel tersebut memungkinkan terdapat dependensi, sehingga diperlukan pemrosesan data (reduksi dimensi) dalam penyusunan. Salah satu metode yang digunakan adalah *Principal Component Analysis* (PCA). Beberapa penelitian untuk pemrosesan data NWP, diantaranya: Bremen (2007); Safitri (2012); Ayuni (2013); Kusumawardani (2014). Bremen (2007) juga menyatakan bahwa reduksi dimensi dengan PCA dapat menguraikan dependensi variabel pada NWP. Disamping itu, Ayuni (2013) menyatakan bahwa reduksi dimensi data NWP dengan PCA menghasilkan nilai RMSEP lebih kecil daripada *kernel Slice Inverse Regression* (kSIR). Oleh karena itu, penelitian ini menggunakan PCA untuk mereduksi dimensi variabel NWP. Sementara itu penelitian tentang penyusunan model MOS telah dilakukan oleh diantaranya Safitri (2012) menggunakan *Projection Pursuit Regression* (PPR). Hasil penelitian tersebut menyimpulkan bahwa hasil dugaan MOS mampu memperbaiki akurasi model NWP. Kusumawardani (2014) menggunakan *Multivariate Regression* dan menghasilkan perbaikan NWP mencapai 90 %. Setiarnoputera (2015) menggunakan *Multivariate Partial Least Squares* (M-PLS) dan menyimpulkan bahwa model PLS2 dapat memperbaiki bias prakiraan NWP.

Berdasarkan penelitian sebelumnya untuk prakiraan cuaca jangka pendek menggunakan MOS masih perlu dikembangkan untuk meningkatkan akurasi prakiraan. Metode yang akan digunakan adalah *Classification and Regression Tree* (CART). Metode CART digunakan untuk menggambarkan hubungan

antara variabel respon dan satu atau lebih variabel prediktor. CART mempunyai kelebihan dibandingkan dengan metode klasifikasi lainnya yaitu hasilnya lebih mudah diinterpretasikan, lebih akurat dan lebih cepat perhitungannya, selain itu CART juga bisa diterapkan untuk himpunan data yang mempunyai jumlah besar dan variabel yang banyak. Akan tetapi, CART juga memiliki kelemahan yaitu menghasilkan pohon yang kurang stabil karena CART sangat sensitif dengan data baru dan bergantung dengan jumlah sampel. Azmi (2017) menyatakan dengan menggunakan CART diperoleh pohon klasifikasi yang layak untuk klasifikasi curah hujan. Metode yang lebih baru dikembangkan untuk klasifikasi adalah *Random Forests* oleh Breiman (2001). Metode ini berusaha menangani beberapa kelemahan yang ada pada metode-metode sebelumnya seperti CART. Beberapa penelitian yang sudah mengaplikasikan *Random Forests* yaitu Dewi (2011) yang menyatakan bahwa *Random Forests* mampu memberikan akurasi yang tinggi dan stabil, Peter (2014) yang menyatakan bahwa *Random Forests* adalah model yang mampu meningkatkan hasil dari kebaikan model dan Vaxjo (2014) yang menyatakan bahwa ketepatan prediksi *Random Forests* lebih baik. Oleh karena itu, pada penelitian ini akan menggunakan CART dan *Random Forests* untuk menyusun model MOS yang akan digunakan untuk prakiraan cuaca jangka pendek.

## **1.2 Rumusan Masalah**

Cuaca memiliki pengaruh yang sangat penting terhadap aktivitas manusia. Seperti halnya informasi cuaca pendek yang sangat dibutuhkan secara cepat dan akurat. Salah satu unsur cuaca pendek yang berperan penting yaitu curah hujan. BMKG telah melakukan penelitian untuk informasi cuaca jangka pendek dengan menggunakan NWP. Namun NWP menghasilkan ramalan yang bias sehingga perlu dilakukan pemrosesan secara statistik menggunakan metode MOS. Pendekatan yang digunakan dalam penyusunan MOS yaitu CART dan *random forests*.

Untuk menyikapi hal tersebut, maka permasalahan yang menjadi fokus dalam penelitian ini adalah sebagai berikut

1. Bagaimana hasil reduksi dimensi variabel NWP dengan menggunakan metode PCA?
2. Bagaimana hasil ketepatan klasifikasi prakiraan curah hujan harian dengan metode CART dan *random forests*?
3. Bagaimana kinerja model hasil ketepatan klasifikasi antara metode CART dan *random forests*?

### 1.3 Tujuan

Berdasarkan rumusan masalah yang telah dipaparkan, maka tujuan penelitian ini adalah sebagai berikut

1. Mendapatkan hasil reduksi dimensi variabel NWP dengan menggunakan metode PCA
2. Mengetahui hasil ketepatan klasifikasi prakiraan curah hujan harian dengan metode CART dan *random forests*
3. Mengetahui kinerja model hasil ketepatan klasifikasi antara metode CART dan *random forests*

### 1.4 Manfaat

Manfaat penelitian ini adalah dapat mengaplikasikan metode CART dan *Random Forests* dalam bidang meteorologi, khususnya mengembangkan model peramalan cuaca jangka pendek dengan metode MOS.

### 1.5 Batasan Masalah

Data NWP yang digunakan dalam penelitian ini adalah hasil model WRF-EMS (*Weather Research and Forecasting-Environment Modeling System*) dan unsur cuaca permukaan yang digunakan adalah curah hujan harian.



## BAB 2 TINJAUAN PUSTAKA

### 2.1 *Principal Component Analysis*

*Principal Component Analysis* (PCA) adalah suatu metode yang berhubungan dengan struktur varians-kovarians dari sekelompok variabel dengan menggunakan kombinasi linier. Metode ini memiliki dua tujuan umum, yaitu untuk mereduksi data dan menginterpretasikan data (Johnson & Wichern, 2007). Prinsip utama metode ini adalah mengelompokkan  $q$  variabel yang saling berkorelasi menjadi  $k$  komponen yang dinamakan komponen utama. Reduksi variabel ini tidak menghilangkan karakteristik data aslinya. Asumsi bahwa data berdistribusi norma multivariat tidak harus dipenuhi dalam metode ini, namun dengan adanya data yang berdistribusi normal multivariat akan mempermudah untuk interpretasi.

Secara umum, vektor random  $\mathbf{x}^T = [x_1, x_2, \dots, x_p]$  memiliki matriks varians-kovarians  $\Sigma$  dengan akar ciri (eigen value) berturut-turut yaitu  $\lambda_1 \geq \lambda_2 \geq \dots \lambda_p \geq 0$ , sehingga kombinasi linier utama seperti pada persamaan berikut (Johnson & Wichern, 2007).

$$\begin{aligned} Y_1 &= \mathbf{e}_1^T \mathbf{x} = e_{11}x_1 + e_{12}x_2 + \dots + e_{1p}x_p \\ Y_2 &= \mathbf{e}_2^T \mathbf{x} = e_{21}x_1 + e_{22}x_2 + \dots + e_{2p}x_p \\ &\vdots \\ Y_p &= \mathbf{e}_p^T \mathbf{x} = e_{p1}x_1 + e_{p2}x_2 + \dots + e_{pp}x_p \end{aligned} \quad (2.1)$$

Jadi, kombinasi linier ke  $-i$  dinyatakan pada persamaan bersamaan  $Y_i = \mathbf{e}_i^T \mathbf{x}$  dimana  $i = 1, 2, \dots, p$  dengan identitas  $Y_i$  seperti ditunjukkan pada persamaan berikut

$$\text{Var}(Y_i) = \mathbf{e}_i^T \Sigma \mathbf{e}_i = \lambda_i, \quad i = 1, 2, \dots, p \quad (2.2)$$

$$\text{Cov}(Y_i, Y_m) = \mathbf{e}_i^T \Sigma \mathbf{e}_m = 0, \quad i \neq m \quad (2.3)$$

PC diharapkan tidak saling berkorelasi dan memiliki varians yang sama dengan akar ciri (*eigen value*) dari  $\Sigma$ , sehingga total varians (keragaman) dapat dijelaskan oleh rumus berikut

$$\sum_{i=1}^p \text{Var}(x_i) = \sigma_{11} + \sigma_{22} + \dots + \sigma_{pp} = \sum_{i=1}^p \text{Var}(Y_i) = \lambda_1 + \lambda_2 + \dots + \lambda_p \quad (2.4)$$

Akibatnya, proporsi varians populasi yang dapat diterangkan oleh PC ke- $i$  ditunjukkan pada persamaan berikut

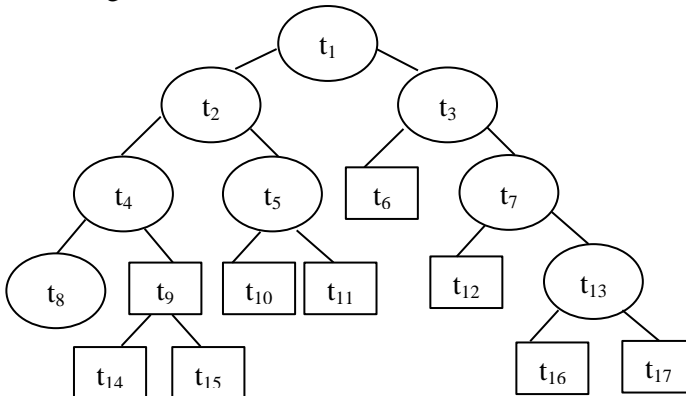
$$\text{Proporsi varians } PC_i = \frac{\lambda_i}{\lambda_1 + \lambda_2 + \dots + \lambda_p} \quad (2.5)$$

Apabila PC yang diambil sebanyak  $m$  dimana ( $m < p$ ), maka persamaan menjadi

$$\text{Proporsi varians } m \text{ } PC = \frac{\lambda_1 + \lambda_2 + \dots + \lambda_m}{\lambda_1 + \lambda_2 + \dots + \lambda_p} \quad (2.6)$$

## 2.2 Classification and Regression Trees

*Classification and Regression Trees* (CART) adalah pendekatan model nonparametrik yang dapat menjelaskan variabel respon yang dipengaruhi oleh variabel prediktor yang bersifat kontinu maupun kategorik. Data *dependent* tergantung dari partisi serangkaian *node* yang bercabang ke kanan dan ke kiri dapat disebut simpul anak (*child nodes*) yang berasal dari simpul utama (*parent node*). Setelah partisi telah berhenti, *child nodes* disebut sebagai *terminal nodes*.



**Gambar 2.1** Ilustrasi Pohon Klasifikasi

Ilustrasi pohon klasifikasi ditunjukkan pada Gambar 2.1. Simpul awal yang merupakan variabel terpenting dalam menduga kelas amatan disebut sebagai simpul utama (*parent node*) dengan notasi  $t_1$ . Simpul dalam (*internal nodes*) dinotasikan dengan  $t_2, t_3, t_4, t_5, t_7, t_9$  dan  $t_{13}$ , serta simpul akhir yang disebut sebagai simpul akhir (*terminal nodes*) dinotasikan dengan  $t_6, t_8, t_{10}, t_{11}, t_{12}, t_{14}, t_{15}, t_{16}$  dan  $t_{17}$  dimana setelahnya tidak ada lagi pemilihan. Setiap simpul berada pada kedalaman (*depth*) tertentu dimana  $t_1$  berada kedalaman 1,  $t_2$  dan  $t_3$  berada pada kedalaman 2, dan begitu seterusnya hingga  $t_{14}, t_{15}, t_{16}$  dan  $t_{17}$  yang berada pada kedalaman 5.

### 2.2.1 Pembentukan Pohon Klasifikasi

Proses pembentukan pohon klasifikasi terdiri atas 3 tahapan, yaitu:

#### 1. Pemilihan (*Classifier*)

Data yang digunakan pada tahap ini adalah sampel data *training/learning* ( $L$ ) yang kemudian dipilah berdasarkan aturan pemilihan dan kriteria *goodness of split*. Himpunan bagian yang dihasilkan dari proses pemilihan harus lebih homogen dibandingkan simpul induksnya. Hal ini dapat dilakukan dengan mendefinisikan fungsi keheterogenitas simpul (*impurity atau  $i(t^*)$* ). Fungsi heterogenitas yang umum digunakan adalah Indeks Gini. Metode ini memiliki kelebihan yaitu proses perhitungan yang sederhana dan relatif cepat, serta mudah dan sesuai untuk diterapkan dalam berbagai kasus (Breiman, Friedman, Olshen, & Stone, 1993). Fungsi Indeks Gini dituliskan dalam persamaan (2.7)

$$i(t)^* = \sum_{j=1}^J p(j|t)p(i|t), i \neq j \quad (2.7)$$

dengan  $p(j|t)$  adalah proporsi kelas  $j$  pada simpul  $t$  dan  $p(i|t)$  adalah proporsi kelas  $i$  pada simpul  $t$ . Setelah dilakukan pemilihan dari semua kemungkinan pemilah, maka tahapan berikutnya adalah menentukan kriteria *goodness of split* ( $\phi(s, t)$ ) untuk mengevaluasi pemilah dari pemilah  $s$  pada simpul  $t$ .

*Goodness of split* ( $\phi(s, t)$ ) didefinisikan sebagai penurunan heterogenitas sebagai berikut.

$$\phi(s, t) = \Delta i(s, t) = i(t) - p_L i(t_L) - p_R i(t_R) \quad (2.8)$$

dengan

- $i(t)$  = fungsi heterogenitas pada simpul  $t$
- $p_L$  = proporsi pengamatan simpul kiri
- $p_R$  = proporsi pengamatan menuju simpul kanan
- $i(t_L)$  = fungsi heterogenitas pada simpul anak kiri
- $i(t_R)$  = fungsi heterogenitas pada simpul anak kanan

$$\Delta i(s^*, t_1) = \max_{s=S} \Delta i(s, t) \quad (2.9)$$

Pemilah yang menghasilkan  $\phi(s, t)$  lebih tinggi merupakan pemilah terbaik karena mampu mereduksi heterogenitas lebih tinggi. Pengembangan pohon dilakukan dengan pencarian pemilah yang mungkin pada simpul  $t_1$  yang kemudian akan dipilah menjadi  $t_2$  dan  $t_3$  oleh pemilah  $s^*$  dan begitu seterusnya.

## 2. Penentuan Simpul Terminal

Suatu simpul  $t$  akan menjadi simpul terminal atau tidak, akan dipilih kembali bila pada simpul  $t$  tidak terdapat penurunan keheterogenan secara berarti atau adanya batasan minimum  $n$  seperti halnya terdapat satu pengamatan pada tiap simpul anak. Jumlah minimum dalam suatu terminal akhir umumnya adalah 5, dan apabila hal itu terpenuhi maka pengembangan pohon dihentikan (Breiman, Friedman, Olshen, & Stone, 1993).

## 3. Penandaan Label Kelas

Penandaan label kelas pada *terminal nodes* dilakukan berdasarkan aturan jumlah terbanyak. Label kelas simpul terminal  $t$  adalah  $j_0$  yang memberi nilai dugaan kesalahan pengklasifikasian simpul  $t$  terbesar. Proses pembentukan pohon klasifikasi berhenti saat terdapat hanya satu pengamatan dalam tiap-tiap simpul anak atau adanya batasan minimum  $n$ , semua pengamatan dalam tiap simpul anak identik dan adanya batasan jumlah level/kedalaman pohon maksimal.

$$p(j_0 | t) = \max_j p(j | t) = \max_j \frac{N_j(t)}{N(t)} \quad (2.10)$$

dengan  $N_j(t)$  merupakan banyaknya amatan kelas  $j$  pada *terminal nodes*  $t$ , dan  $N(t)$  merupakan jumlah total pengamatan dalam *terminal node*  $t$ . Label kelas untuk *terminal node*  $t$  adalah  $j_0$  yang memberikan nilai dugaan kesalahan pengklasifikasian pada simpul  $t$  paling kecil sebesar  $r(t) = 1 - \max_j p(j|t)$ .

### 2.2.2 Pemangkasan Pohon Klasifikasi

Bagian pohon yang kurang penting dilakukan pemangkasan sehingga didapatkan pohon klasifikasi yang optimal. Pemangkasan didasarkan pada suatu penilaian ukuran sebuah pohon tanpa mengorbankan kebaikan ketepatan melalui pengurangan simpul pohon sehingga dicapai ukuran pohon yang layak. Ukuran pemangkasan yang digunakan untuk memperoleh ukuran pohon yang layak tersebut adalah *cost complexity minimum* (Lewis, 2000).

$$R_\alpha(T) = R(T) + \alpha |\bar{T}| \quad (2.11)$$

dimana

$R(T)$  = *resubtitution estimate* (proporsi kesalahan pada subpohon)

$\alpha$  = *complexity parameter* (kompleksitas parameter)

$|\bar{T}|$  = ukuran banyaknya simpul terminal pohon  $T$

### 2.2.3 Penentuan Pohon Klasifikasi Optimal

Ukuran pohon yang terlalu besar akan menyebabkan nilai *cost complexity* yang tinggi karena struktur data yang digambarkan cenderung kompleks sehingga perlu dipilih pohon optimal yang berukuran sederhana tetapi memberikan nilai penduga pengganti yang cukup kecil. Bila  $R(T)$  dipilih sebagai penduga terbaik, maka akan cenderung dipilih pohon yang besar, sebab pohon yang semakin besar akan membuat nilai  $R(T)$  semakin kecil. Terdapat dua macam penduga untuk mendapatkan pohon klasifikasi optimal yaitu penduga sampel uji (*test sample*

*estimate*) dan penduga validasi silang lipat  $v$  (*cross validation v-fold estimate*).

a. Penduga Sampel Uji (*Test Sample Estimate*)

Penduga sampel uji digunakan ketika data berukuran besar. Prosedur *test sample estimate* diawali dengan membagi data *learning* menjadi dua bagian yaitu  $L_1$  dan  $L_2$ . Pengamatan dalam  $L_1$  digunakan untuk membentuk pohon  $T$ , sedangkan pengamatan dalam  $L_2$  digunakan untuk menduga  $R^{ts}(T)$ . Persamaan *test sample estimate* adalah sebagai berikut.

$$R^{ts}(T) = \frac{1}{N_2} \sum_{(x_n, j_n) \in L_2} X(d(x_n) \neq j_n) \quad (2.12)$$

dengan  $N_2$  adalah jumlah pengamatan dalam  $L_2$  dan  $X(\cdot)$  bernilai 0 jika pertanyaan dalam tanda kurung salah dan bernilai 1 jika pertanyaan dalam tanda kurung benar. Pohon klasifikasi yang optimum dipilih  $T^*$  dengan  $R^{ts}(T^*) = \min R^{ts}(T)$ .

b. Penduga Validasi Silang Lipat  $V$  (*Cross Validation V-Fold Estimate*)

Penduga pengganti ini sering digunakan apabila pengamatan yang ada tidak cukup besar. Pengamatan dalam  $L$  dibagi secara random menjadi  $V$  bagian yang saling lepas dengan ukuran kurang lebih sama besar untuk setiap kelas. Pohon  $T^{(v)}$  dibentuk dari sampel *learning* ke  $v$  dengan  $v = 1, 2, \dots, V$ . Dimisalkan  $d^{(v)}(x)$  adalah hasil pengklasifikasian, maka penduga sampel uji untuk  $R(T^{(v)})$  adalah sebagai berikut.

$$R(T^{(v)}) = \frac{1}{N_v} \sum_{(x_n, j_n) \in L_v} X(d^{(v)}(x_n) \neq j_n) \quad (2.13)$$

dengan  $N_v \cong N/V$  adalah jumlah pengamatan dalam  $L_v$ .

Selanjutnya, dilakukan prosedur yang sama dengan menggunakan semua pengamatan dalam  $L$  untuk membentuk deret pohon  $T_t$ . Penduga validasi silang lipat  $v$  untuk  $T_t^{(v)}$  adalah

$$R^{cv}(T_t) = \frac{1}{V} \sum_{v=1}^V R^{cv}(T_t^{(v)}) \quad (2.14)$$

Pohon klasifikasi yang optimum dipilih  $T^*$  dengan  $R^{cv}(T^*) = \min_t R^{cv}(T_t)$ .

### 2.3 *Random Forests*

*Random forests* adalah suatu metode klasifikasi yang terdiri dari gabungan pohon klasifikasi (CART) yang saling independen yang berasal dari distribusi yang sama melalui proses *voting* (jumlah terbanyak) untuk memperoleh prediksi klasifikasi. *Random forests* merupakan pengembangan dari metode *ensemble* yang pertama kali dikembangkan oleh Leo Breiman (2001) yang digunakan untuk meningkatkan ketepatan klasifikasi. Bila dalam proses *bagging* digunakan *resampling bootstrap* untuk membangkitkan pohon klasifikasi dengan banyak versi yang kemudian mengkombinasikannya untuk memperoleh prediksi akhir, maka dalam *random forests* proses pengacakan untuk membentuk pohon klasifikasi tidak hanya dilakukan untuk data sampel saja melainkan juga pada pengambilan variabel prediktor. Sehingga, proses ini akan menghasilkan kumpulan pohon klasifikasi dengan ukuran dan bentuk yang berbeda-beda. Hasil yang diharapkan adalah suatu kumpulan pohon klasifikasi yang memiliki korelasi kecil antar pohon. Korelasi yang kecil akan menurunkan hasil kesalahan prediksi *Random Forests* (Breiman, 2001).

Secara umum, pengembangan *random forests* yang dilakukan dari proses *bagging* yaitu terletak pada proses pemilah. Pada *random forests* pemilihan pemilah hanya melibatkan beberapa variabel prediktor yang diambil secara acak. Algoritma *random forests* dijelaskan sebagai berikut.

- a. Mengambil  $n$  data sampel dari *dataset* awal dengan menggunakan teknik *resampling bootstrap* dengan pengembalian.
- b. Menyusun pohon klasifikasi dari setiap *dataset* hasil *resampling bootstrap*, dengan penentuan pemilah terbaik didasarkan pada variabel prediktor yang diambil secara acak. Jumlah variabel yang diambil secara acak dapat

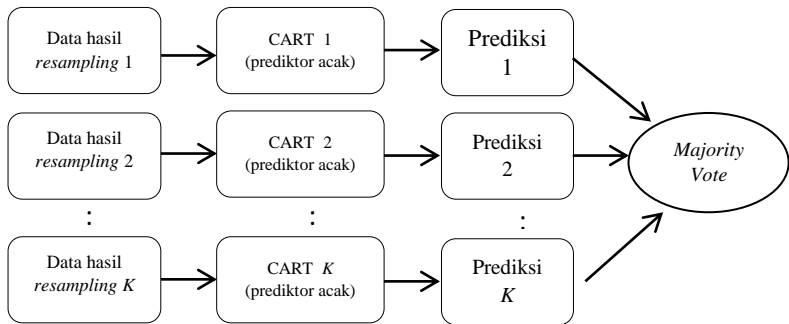
- ditentukan melalui perhitungan  $\log_2(M + 1)$  dimana  $M$  adalah banyak variabel prediktor (Breiman, 2001).
- Melakukan prediksi klasifikasi data sampel berdasarkan pohon klasifikasi yang terbentuk.
  - Mengulangi langkah a-c hingga diperoleh sejumlah pohon klasifikasi yang diinginkan. Perulangan dilakukan sebanyak  $K$  kali.
  - Melakukan prediksi klasifikasi data sampel akhir dengan mengkombinasikan hasil prediksi pohon klasifikasi yang diperoleh berdasarkan aturan *majority vote*.

Dalam analisis dengan menggunakan metode *random forests* dimulai dari pengambilan data dengan teknik *resampling bootstrap*. *Bootstrap* adalah suatu metode yang dapat bekerja tanpa membutuhkan asumsi distribusi karena sampel asli digunakan sebagai populasi. *Bootstrap* pertama kali diperkenalkan oleh Efron (1979) yang digunakan untuk mencari distribusi sampling dari suatu estimator dengan prosedur *resampling* dengan pengembalian dari data asli (Sungkono, 2013). Berikut adalah algoritma dari *resampling bootstrap*

- Mengkonstruksi distribusi empiris  $\hat{F}_n$  dari suatu sampel dengan memberikan probabilitas  $1/n$  pada setiap  $X_i$  dimana  $i=1,2, \dots, n$ .
- Mengambil sampel *bootstrap* berukuran  $n$  secara random dengan pengembalian dari distribusi empiris  $\hat{F}_n$  disebut sebagai sampel *bootstrap* pertama  $X^{*1}$ .
- Menghitung statistik  $\hat{\theta}$  yang diinginkan dari sampel *bootstrap*  $X^{*1}$  disebut sebagai  $\hat{\theta}_1^*$ .
- Mengulangi langkah b dan c hingga  $B$  kali diperoleh  $\hat{\theta}_1^*, \hat{\theta}_2^*, \dots, \hat{\theta}_B^*$ .



- e. Mengkontruksi suatu distribusi probabilitas dari  $\hat{\theta}_B^*$  dengan memberikan probabilitas  $1/B$  pada setiap  $\hat{\theta}_1^*$ ,  $\hat{\theta}_2^*$ , ...,  $\hat{\theta}_B^*$ . Distribusi tersebut merupakan estimator *bootstrap* untuk distribusi sampling  $\hat{\theta}$  dan dinotasikan dengan  $\hat{F}^*$
- f. Pendekatan estimasi *bootstrap* adalah  $\hat{\theta}^* = \sum_{b=1}^B \hat{\theta}_b^* \frac{1}{B}$



**Gambar 2.2** Algoritma *Random Forests*

### 2.3 Ukuran Ketepatan Klasifikasi

Ketepatan klasifikasi digunakan untuk mengetahui apakah data diklasifikasikan dengan benar atau tidak. Ketepatan klasifikasi merupakan suatu evaluasi untuk melihat peluang kesalahan klasifikasi yang dilakukan oleh suatu fungsi klasifikasi. Beberapa cara yang umum digunakan untuk mengukur ketepatan klasifikasi adalah menggunakan *apparent error rate* (APER) dan *total accuracy rate* (1-APER). APER merupakan proporsi observasi yang diprediksi secara tidak benar (ukuran kesalahan klasifikasi total). *Total accuracy rate* merupakan proporsi observasi yang diprediksi secara benar (ukuran kesalahan klasifikasi total) (Johnson & Wichern, 2007).

Tabel untuk menghitung ketepatan klasifikasi ditunjukkan pada tabel berikut

**Tabel 2.1** *Crosstab* Ketepatan Klasifikasi

Aktual Variabel Y	Prediksi Variabel Y				Total
	1	2	..	k	
1	$n_{11}$	$n_{12}$	..	$n_{1k}$	$N_{1.}$
2	$n_{21}$	$n_{22}$	..	$n_{2k}$	$N_{2.}$
:	:	:	..	:	:
k	$n_{k1}$	$n_{k2}$	..	$n_{kk}$	$N_{k.}$
Total	$N_{.1}$	$N_{.2}$	..	$N_{.k}$	$N$

Dengan:

$i = 1, 2, \dots, k$

$n_{ii}$  : jumlah subjek dari variabel Y kategori  $i$  yang tepat diprediksikan sebagai variabel Y kategori  $i$

$N_{i.}$  : jumlah observasi dari variabel Y kategori  $i$

$N_{.i}$  : jumlah prediksi dari variabel Y kategori  $i$

$N$  : jumlah total observasi dan prediksi

Adapun rumus yang digunakan sebagai berikut

$$\text{Total accuracy rate} = 1 - \text{APER} = \frac{n_{11} + n_{22} + \dots + n_{kk}}{N} \quad (2.15)$$

$$\text{APER} = 1 - \text{Total accuracy rate} \quad (2.16)$$

## 2.4 Model Output Statistics

Model NWP adalah sekumpulan kode komputer yang mempresentasikan secara numerik persamaan-persamaan atmosfer, digunakan untuk memprediksi kondisi atau status atmosfer yang akan datang dengan menggunakan kemampuan komputer yang tinggi. NWP merupakan grid dengan resolusi 7-60 km. Hasil ramalan NWP untuk lokasi tertentu dengan resolusi tinggi seringkali bias terutama lokasi dengan topografi dan vegetasi yang kompleks, karena *forcing* lokal lebih dominan. Untuk mengoptimalkan pemanfaatan keluaran model NWP perlu dilakukan pemrosesan (*post processing*). Beberapa metode yang seringkali digunakan pemrosesan NWP adalah salah satunya MOS.

MOS pertama kali diperkenalkan oleh Glahn dan Lowry pada tahun 1972. MOS merupakan pemodelan yang menentukan hubungan antara variabel respon dan variabel prediktor dari model numerik dalam beberapa ramalan waktu (Glahn & Lowry, 1972). Variabel respon yang digunakan adalah hasil observasi cuaca yang merupakan unsur-unsur cuaca seperti suhu minimum, suhu maksimum, kelembapan minimum, kelembapan maksimum dan curah hujan. Sedangkan variabel prediktornya merupakan parameter yang didapatkan dari luaran NWP seperti temperatur, angin dan sebagainya pada berbagai grid dan level. Disamping itu, variabel prediktor dapat juga berupa parameter geografi seperti lintang, bujur dan waktu ( $t$ ). Tipe data variabel prediktor dapat digunakan beberapa tipe, diantaranya kontinu (seperti temperatur), diskret (seperti klasifikasi awan: 0=cerah, 1=tersebar (*scattered*), 2=patah-patah (*broken*) dan 3=mendung), *binary* (seperti: 1=terjadi presipitasi, 0=tidak terjadi presipitasi) dan diskontinyu.

Pemodelan MOS yang digunakan dalam menentukan hubungan antara variabel respon dan prediktor adalah berbasis regresi. Secara umum persamaan matematis MOS adalah sebagai berikut (Wilks, 2006):

$$\hat{y}_t = f_{MOS}(x_t) \quad (2.17)$$

dimana,

$\hat{y}_t$  = ramalan cuaca saat  $t$

$x_t$  = variabel luaran NWP saat  $t$

Model MOS didapatkan dari hubungan regresi antara data observasi cuaca dengan data luaran NWP pada  $t$  yang sama. Namun dalam implementasinya, ramalan cuaca saat  $(t+1)$  didapatkan berdasarkan data ramalan NWP saat  $(t+1)$  yang telah diketahui pada saat ke- $t$ . Seperti yang telah disebutkan sebelumnya bahwa pemodelan MOS berbasis regresi, sehingga model yang menggambarkan hubungan anatara variabel prediktor ( $X$ ) dengan varaibel respon ( $Y$ ) adalah

$$Y = f(X, \beta) + \varepsilon \quad (2.18)$$

Hubungan antara variabel prediktor dan variabel respon dikatakan linier jika dapat dinyatakan dalam bentuk:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \varepsilon \quad (2.19)$$

Dalam bentuk matriks, model regresi linier berganda dapat ditulis:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2.20)$$

Atau

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & \dots & X_{1p} \\ 1 & X_{21} & \dots & X_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & \dots & X_{np} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Selanjutnya nilai  $\boldsymbol{\beta}$  dapat ditaksir menggunakan metode kuadrat terkecil yaitu:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} (\mathbf{X}^T \mathbf{Y}) \quad (2.21)$$

Ukuran kebaikan model yang digunakan untuk menggambarkan kebaikan prediksi adalah nilai koefisien determinasi ( $R^2$ )

$$R^2 = \frac{\sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (2.22)$$

Secara umum, MOS memiliki dua tujuan utama. Pertama, MOS menghasilkan pendugaan parameter yang mungkin tidak ada pada parameter NWP, seperti peluang hujan, petir, kabut dan lain-lain. Tujuan kedua MOS adalah untuk mengurangi sisaan dari hasil ramalan NWP dengan memperkecil bias dan koreksi model secara statistik. Beberapa syarat agar MOS menghasilkan ramalan yang optimum adalah:

1. Periode data *learning* yang digunakan dalam pembentukan model seharusnya sepanjang mungkin
2. Model yang terbentuk seharusnya tidak berubah pada kondisi ekstrim selama verifikasi model

3. Pada tahap validasi model, MOS seharusnya dapat diaplikasikan dan tidak berubah modelnya. Validasi model dimaksudkan untuk menguji keterandalan model yang sudah terbangun dengan menggunakan data independen.

## 2.5 Numerical Weather Prediction

*Numerical Weather Prediction* (NWP) adalah sekumpulan kode komputer yang mempresentasikan secara numerik persamaan-persamaan atmosfer berdasarkan sifat-sifat fisika dinamis, digunakan untuk memprediksi konsisi atau status NWP atmosfer yang akan datang dengan menggunakan kemampuan komputer yang tinggi. Prediksi cuaca dirumuskan dengan menyelesaikan persamaan pergerakan atmosfer. Terdapat dua keuntungan menggunakan data NWP yaitu data NWP lebih mudah didapatkan ketika terdapat jaringan internet dan model yang diberikan memiliki pola distribusi spasial yang relevan terhadap ukuran area yang diteliti.

NWP diukur dalam domain lokasi atau grid yang tinggi, berkisar antara 7-60 km, dengan skala sebesar itu NWP akan memberikan informasi cuaca yang homogen pada daerah grid tersebut. Kondisi cuaca skala kecil atau skala lokal kurang terepresentasikan dengan baik. Oleh karena itu, luaran NWP juga diukur dengan domain yang tinggi. Selain itu, luaran NWP juga bersifat deterministik sehingga perlu dilakukan pemrosesan secara statistik (Wilks, 2006).

Penelitian ini menggunakan model NWP *Weather Research and Forecasting – Environment Modeling System* (WRF-EMS). Model WRF-EMS merupakan model yang dikembangkan oleh *National Oceanic and Atmospheric Administration* (NOAA) bekerja sama dengan *National Weather Service* (NWS), *Forecast Decision Training Branch* (FDTB), dan *Science Operations Officer and Training Resource Center* (SOO/STCR). Untuk keperluan prakiraan cuaca model WRF-EMS dinilai sebagai model yang lengkap dan merupakan *state of the science* dari produk NWP serta merupakan gabungan dari dua model dinamis

yang umum digunakan di dunia pemodelan yaitu *Advanced Research WRF (ARW)* yang dikembangkan oleh *National Center for Atmospheric Research (NCAR)* dan *non-Hydrostatic Mesoscale Model (NMM)* yang dikembangkan oleh *National Center for Environmental Prediction (NCEP)*. Di dalam WRF-EMS, hampir semua sistem operasional NWP sudah diintegrasikan termasuk proses pengambilan dan pengolahan data, eksekusi model, pengolahan data output serta migrasi hasil output dan penyimpanan, perangkat lunak untuk menampilkan hasil prakiraan juga terdapat dalam model WRF-EMS tersebut (Hutapea, Kurniawan, & Hanggoro, 2016).

## 2.6 Curah Hujan

Curah Hujan (CH) menurut BMKG adalah ketinggian air hujan yang terkumpul dalam penakar hujan pada tempat yang datar, tidak menyerap, tidak meresap dan tidak mengalir. Unsur hujan 1 (satu) milimeter artinya dalam luasan satu meter persegi pada tempat yang datar tertampung air hujan setinggi satu milimeter atau tertampung air hujan sebanyak satu liter. Berdasarkan intensitasnya, curah hujan diklasifikasikan menjadi 5 yakni cerah berawan, hujan ringan, hujan sedang, hujan lebat dan hujan lebat sekali dengan batasan seperti pada Tabel 2.2 berikut.

**Tabel 2.2** Klasifikasi Intensitas Curah Hujan

Klasifikasi Hujan	Intensitas Curah Hujan (mm/hari)
Cerah berawan	$\leq 0.1$
Hujan ringan	$0.1 < CH \leq 20$
Hujan sedang	$20 < CH \leq 50$
Hujan lebat	$50 < CH \leq 100$
Hujan lebat sekali	$> 100$

(Sumber: BMKG, 2006)

## **BAB III**

### **METODOLOGI PENELITIAN**

#### **3.1 Sumber Data**

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari BMKG Pusat Jakarta, yaitu data observasi cuaca permukaan dan data NWP model WRF-EMS yang diunduh dari website NCEP-NOAA yaitu <http://dss.ucar.edu> pada periode 2 Januari 2015 sampai 10 Desember 2016. Lokasi pengamatan yang dilakukan dalam penelitian ini ada 3 stasiun meteorologi di Indonesia yaitu Stasiun Meteorologi Juanda, Stasiun Meteorologi Ngurah Rai dan Stasiun Meteorologi Soetta dengan lintang bujur pada Tabel 3.1. Ketiga stasiun pengamatan tersebut dipilih karena umumnya pengukuran cuaca pada stasiun tersebut dilakukan rutin sehingga data pengamatan yang dihasilkan cukup baik dan lengkap.

**Tabel 3.1** Wilayah Stasiun Pengamatan

No	Kab./Kota	Nama Stasiun	Lintang	Bujur
1	Surabaya	Stasiun Meteorologi Juanda	-7.38	112.78
2	Denpasar	Stasiun Meteorologi Ngurah Rai	-8.75	115.16
3	Jakarta	Stasiun Meteorologi Soetta	-6.12	106.65

#### **3.2 Variabel Penelitian**

Variabel penelitian yang digunakan adalah hasil observasi cuaca permukaan sebagai variabel respon dan data model *Numerical Weather Prediction* (NWP) sebagai variabel prediktor.

Variabel respon yang digunakan adalah curah hujan harian (mm) yang diperoleh dari hasil pengukuran di 3 stasiun pengamatan yaitu Stasiun Meteorologi Juanda, Ngurah Rai dan Soetta. Untuk variabel curah hujan akan dilakukan transformasi data menjadi klasifikasi 5 tipe hujan, dengan kriteria sebagai berikut

**Tabel 3.2** Klasifikasi Curah Hujan Menurut Intensitasnya

Klasifikasi Hujan	Intensitas Curah Hujan (mm/hari)
Cerah berawan	Curah hujan $\leq 0.1$
Hujan ringan	$0.1 < \text{Curah hujan} \leq 20$
Hujan sedang	$20 < \text{Curah hujan} \leq 50$
Hujan lebat	$50 < \text{Curah hujan} \leq 100$
Hujan lebat sekali	Curah hujan $> 100$

(Sumber: BMKG, 2006)

Variabel prediktor yang digunakan adalah data NWP model WRF-EMS (*Weather Research and Forecasting – Environment Modeling System*) yang dikembangkan oleh *National Oceanic and Atmospheric Administration (NOAA)*. Variabel yang akan digunakan dari NWP WRF-EMS sebanyak 15 variabel seperti yang disajikan dalam Tabel 3.3

**Tabel 3.3** Variabel NWP WRF-EMS

No	Nama Variabel	Satuan
1	Hujan Total (hujan total)	mm/jam
2	<i>Dew Point 2m</i> (dew2m)	celcius
3	<i>Wind Gust</i> (wind gust)	m/s
4	<i>High Cloud</i> (hi cloud)	%
5	<i>Low Cloud</i> (low cloud)	%
6	<i>Middle Cloud</i> (mid cloud)	%
7	<i>Total Cloud</i> (tot cloud)	%
8	<i>Mean Sea Level Pressure</i> (mslp)	mb
9	<i>Surface Pressure</i> (sur press)	mb
10	<i>Relative Humidity 2m</i> (rh2m)	%
11	<i>Specific Humidity 2m</i> (sh2m)	kg/kg
12	<i>Temperature 2m</i> (temp2m)	celcius
13	<i>U Wind 10m</i> (uwind10m)	m/s
14	<i>Vwind 10m</i> (vwind10m)	m/s
15	<i>Surface Vis</i> (sur vis)	meter

Berikut merupakan definisi dari masing-masing variabel NWP yang digunakan sebagai variabel prediktor

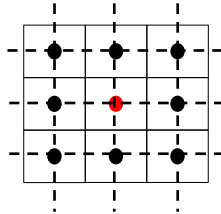
1. Hujan Total merupakan potensi hujan
2. *Dew Point 2 m* merupakan pengukuran titik embun yang dilakukan pada ketinggian 2 meter.
3. *Wind Gust* merupakan angin yang berkaitan dengan badai guntur arah dan kecepatannya dapat berubah dengan cepat



pada waktu terik matahari. Perubahan kecepatannya dapat besar dan dalam waktu yang singkat. Sesaat sesudah terjadi perubahan angin tersebut, biasanya diikuti keadaan lengang atau tak ada angin.

4. *High Cloud* merupakan awan yang letaknya pada ketinggian lebih dari 3 km tetapi tidak sama disetiap wilayah.
5. *Low Cloud* merupakan awan yang terletak pada paras rendah kurang dari 2 km.
6. *Middle Cloud* merupakan awan yang terletak pada paras antara 2 dan 8 km tetapi tidak sama disetiap wilayah
7. *Total Cloud* merupakan jumlah awan yang terletak pada paras rendah dan tinggi.
8. *Mean Sea Level Pressure* merupakan suhu rata-rata di atas permukaan laut
9. *Surface pressure* merupakan tekanan permukaan diukur berdasarkan tekanan gaya pada permukaan dengan luas tertentu. Tekanan udara berkurang dengan bertambahnya ketinggian.
10. *Relative humidity 2m* merupakan nilai perbandingan antara tekanan uap air yang ada pada saat pengukuran dan besarnya tekanan uap air maksimum yang dapat dicapai pada suhu dan tekanan udara pada saat itu yang di ukur pada 2 meter di atas permukaan laut.
11. *Specific humidity* merupakan perbandingan antara massa uap air dan massa udara dalam tiap satuan volume udara yang di ukur pada 2 meter di atas permukaan laut.
12. *Temperatur 2 m* merupakan derajat panas dan dingin udara di atmosfer yang diukur pada ketinggian 2 meter di atas permukaan.
13. *U wind 10m* merupakan komponen U angin yang bergerak dengan arah barat-timur pada ketinggian 10 meter diatas permukaan.
14. *V wind 10m* merupakan komponen V angin yang bergerak dengan arah utara-selatan pada ketinggian 10 meter diatas permukaan.

Semua variabel NWP WRF-EMS ini akan diukur pada 9 grid pengukuran terdekat dari lokasi stasiun pengamatan sehingga untuk domain 3x3 terdapat 9 variabel prediktor. Oleh karena itu jumlah variabel yang digunakan dalam pemodelan adalah 135 variabel prediktor. Resolusi grid yang digunakan adalah  $0,5^0$ . Proyeksi pengukuran variabel NWP dalam grid 3x3 ditunjukkan pada Gambar 3.1



**Gambar 3.1** Pengukuran NWP dalam Grid 3x3

Titik berwarna merah pada Gambar 3.1 menunjukkan grid terdekat pada lokasi stasiun pengamatan, sedangkan titik berwarna hitam merupakan kombinasi grid di sekitar lokasi pengamatan. Oleh karena itu, masing-masing variabel akan mempunyai korelasi yang kuat dengan dirinya sendiri karena diukur pada 9 grid pengukuran. Sedangkan antar variabel prediktor juga akan mempunyai korelasi yang kuat, hal ini dikarenakan data NWP berdimensi tinggi. Karena memiliki dimensi yang cukup besar, maka perlu dilakukan reduksi dimensi pada grid pengukuran variabel NWP yaitu salah satunya menggunakan *Principal Component Analysis* (PCA).

### 3.3 Struktur Data

Struktur data pada penelitian ini disajikan dalam Tabel 3.4 berikut.

**Tabel 3.4** Struktur Data Penelitian

Hari ke-	Y	X <sub>1</sub>				X <sub>15</sub>		
		Grid 1	...	Grid 9	...	Grid 1	...	Grid 9
1	Y <sub>1</sub>	X <sub>1(1)</sub>	...	X <sub>1(9)</sub>	...	X <sub>1(1)</sub>	...	X <sub>1(9)</sub>
:	:	:	...	:	...	:	...	:
:	:	:	...	:	...	:	...	:
353	Y <sub>1</sub>	X <sub>353(1)</sub>	...	X <sub>353(9)</sub>	...	X <sub>353(1)</sub>	...	X <sub>353(9)</sub>

Keterangan :

Y = curah hujan harian

X = variabel NWP

X<sub>1(1)</sub> = variabel NWP pada hari ke-1 untuk grid 1

### 3.4 Langkah Analisis

Langkah-langkah analisis data yang dilakukan dalam penelitian ini adalah sebagai berikut

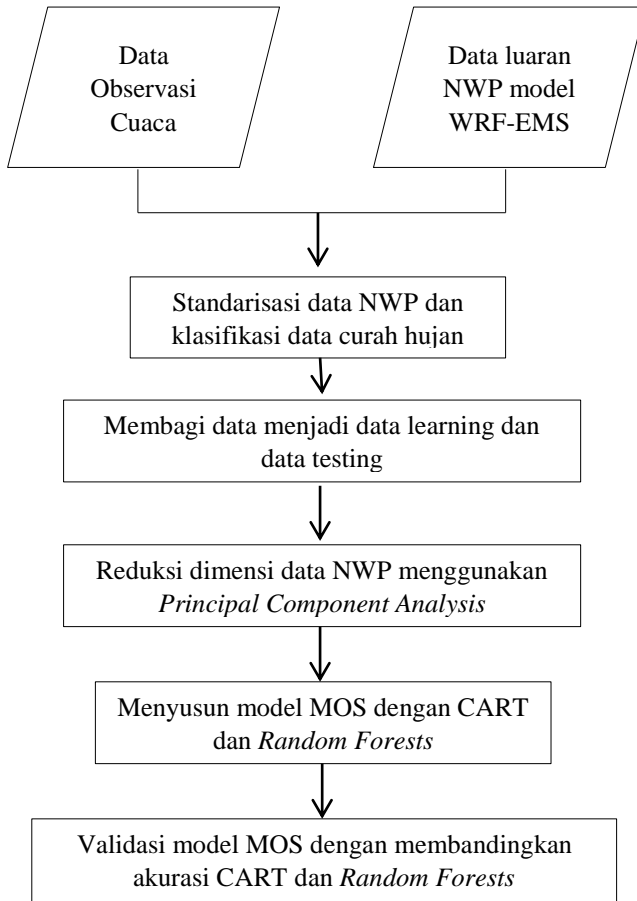
1. Melakukan pra-pemrosesan data secara statistik
  - a. Mengecek *missing value* pada data curah hujan, jika ada kode 8888 (tidak terukur) diganti dengan nol dan jika kode 9999 (*missing value*) dilakukan imputasi dengan menghitung rata-rata hari sebelumnya dan hari sesudahnya
  - b. Melakukan standarisasi data output NWP dan mengklasifikasikan data curah hujan sesuai dengan kategori dari BMKG.
  - c. Membagi data menjadi dua bagian yaitu data *learning* dan data *testing*. Data *testing* diambil sebanyak 30 data terbaru dan sisanya dijadikan data *learning*.
2. Melakukan reduksi dimensi dari variabel NWP menggunakan metode *Principal Component Analysis*. Berikut langkah analisis yang dilakukan
  - i) Menghitung matriks varians-covarians dari variabel NWP
  - ii) Menghitung nilai *eigenvalue* dan nilai normalisasi *eigenvector*

- iii) Menghitung proporsi kumulatif varians dari *eigenvalue* yang diperoleh
- iv) Menentukan jumlah komponen utama yang digunakan dengan melihat kumulatif varians yang lebih dari 80%.
- 3. Melakukan pemodelan MOS menggunakan *hybrid* CART dan *Random Forests* pada 3 stasiun pengamatan untuk memprakirakan curah hujan. Berikut langkah analisis yang dilakukan
  - a. Menyusun Model CART
    - i) Membentuk pohon klasifikasi dengan data *learning* melalui langkah pemilihan pemilah dari variabel prediktor berdasarkan aturan pemilahan Indeks Gini, kemudian hasil pemilahan dievaluasi dengan menggunakan kriteria *goodness of split*
    - ii) Menentukan *terminal nodes* dengan cara menghentikan pembentukan pohon hingga mencapai batasan minimum pengamatan dalam terminal *node* dan melakukan penandaan label kelas pada terminal node berdasarkan aturan jumlah terbanyak dari tiap kelas pada variabel respon.
    - iii) Memangkas pohon (*pruning*) berdasarkan kriteria *cost complexity minimum*.
    - iv) Menentukan pohon klasifikasi optimal dengan *10-Fold Cross Validation Estimate*.
    - v) Menghitung ketepatan klasifikasi pohon hasil bentukan menggunakan data *learning* dan mengevaluasinya dengan menjalankan data testing pada pohon klasifikasi yang terbentuk. Ukuran ketepatan klasifikasi yang digunakan yaitu APER dan *total accuracy rate* (1-APER).
  - b. Menyusun Model dengan *Random Forests*
    - i) Mengambil  $n$  sampel *bootstrap* dengan pengembalian dari data *learning*
    - ii) Menentukan jumlah variabel prediktor yang akan dilakukan pengambilan secara acak dalam proses

penentuan pemilah saat pembentukan pohon klasifikasi melalui perhitungan  $\log_2(M + 1)$  dimana  $M$  adalah banyak variabel prediktor.

- iii) Membentuk pohon klasifikasi dimana pemilihan *node* terbaik dilakukan berdasarkan variabel-variabel prediktor yang diambil secara acak
  - iv) Melakukan prediksi klasifikasi untuk data *learning*
  - v) Mengulangi langkah i-iii hingga  $K$  kali replikasi
  - vi) Melakukan *voting* mayoritas hasil prediksi klasifikasi dari  $K$  kali replikasi pembentukan pohon klasifikasi
  - vi) Menghitung ketepatan klasifikasi menggunakan data *learning* dan data *testing* menggunakan kriteria APER dan *total accuracy rate* (1-APER) sebanyak  $K$  kali replikasi.
4. Melakukan validasi model MOS dengan membandingkan tingkat akurasi klasifikasi antara CART dan *Random Forests*.

Langkah analisis yang telah dijabarkan dapat dirangkum dalam sebuah *flowchart* seperti Gambar 3.2 berikut ini



**Gambar 3.2** *Flowchart* Langkah Analisis

## BAB IV ANALISIS DAN PEMBAHASAN

Bab ini membahas penyusunan MOS dengan metode klasifikasi pohon dan *random forest* dengan menghitung ketepatan klasifikasi menggunakan nilai 1-APER. Bagian awal disajikan deskripsi curah hujan di 3 stasiun pengamatan dan membahas reduksi dimensi data NWP menggunakan metode PCA.

### 4.1 Deskripsi Curah Hujan

Deskripsi curah hujan menurut BMKG berdasarkan intensitasnya disajikan pada Tabel 4.1 berikut

**Tabel 4.1** Persentase Kejadian Hujan

Stasiun Pengamatan	Kategori Kejadian Hujan (%)					Total
	Cerah Berawan	Hujan Ringan	Hujan Sedang	Hujan Lebat	Hujan Lebat Sekali	
Juanda	58.36	27.20	11.05	3.40	0.00	100
Ngurah Rai	64.87	29.18	3.68	1.70	0.57	100
Soetta	64.31	28.61	4.82	1.70	0.57	100

Tabel 4.1 menunjukkan bahwa pada periode 2 Januari 2015 hingga 10 Desember 2016 pada Stamet Juanda hujan sedang dan hujan lebat lebih banyak terjadi dibandingkan Stamet lainnya. Namun hujan lebat sekali tidak pernah terjadi di Stamet Juanda. Di Stamet Ngurah Rai, cerah berawan dan hujan ringan memiliki persentase paling tinggi. Sedangkan untuk hujan sedang memiliki persentase paling rendah di Stamet Ngurai Rai. Hujan lebat dan hujan lebat sekali pada Stamet Ngurai Rai dan Soetta memiliki persentase yang sama.

#### 4.2 Reduksi Dimensi Data Numerical Weather Prediction

Kriteria penentuan variabel baru dari reduksi dimensi menggunakan metode PCA yaitu berdasarkan besar proporsi keragaman yang dapat dijelaskan oleh komponen yang terbentuk diatas 80 persen. *Eigenvalue* dan keragaman kumulatif variabel hujan\_tot hasil reduksi PCA pada Stamet Juanda ditampilkan pada Tabel 4.2 berikut

**Tabel 4.2** *Eigenvalue* dan Kumulatif Keragaman Variabel hujan total Stamet Juanda

PC	<i>Eigenvalue</i>	Keragaman yang dijelaskan	Keragaman Kumulatif
1	6.150	0.683	0.683
2	1.558	0.173	0.857
3	0.662	0.074	0.930
4	0.207	0.023	0.953
5	0.188	0.021	0.974
6	0.103	0.011	0.986
7	0.082	0.009	0.995
8	0.031	0.004	0.998
9	0.015	0.002	1.000

Keragaman variabel hujan\_total yang dijelaskan oleh komponen (PC) pertama sebesar 68.3 persen dan pada komponen kedua sebesar 17.3 persen. Keragaman kumulatif pada komponen kedua mencapai 85.7 persen sehingga keragaman variabel hujan\_total dapat dijelaskan dengan 2 komponen. Jumlah kumulatif keragaman diatas 80 persen yang terbentuk dari data NWP secara lengkap disajikan pada Tabel 4.3 berikut

**Tabel 4.3** *Eigenvalue* dan Keragaman PC Stamet Juanda

Variabel	Juanda		
	Jumlah PC	<i>Eigenvalue</i>	Keragaman Kumulatif
hujan_total	2	6.150; 1.558	85.7%
dew2m	1	8.610	95.7%



**Tabel 4.3** *Eigenvalue* dan Keragaman PC Stamet Juanda (Lanjutan)

Variabel	Juanda		
	Jumlah PC	Eigenvalue	Keragaman Kumulatif
wind_gust	1	7.493	83.3%
hi_cloud	1	8.227	91.4%
low_cloud	2	5.640; 1.703	81.6%
mid_cloud	1	8.172	90.8%
tot_cloud	1	8.247	91.6%
mslp	1	8.948	99.4%
sur_press	1	8.918	99.1%
rhum2m	1	7.924	88.0%
shum2m	1	8.610	95.7%
temp2m	2	6.562; 1.950	94.6%
uwind10m	1	8.850	98.3%
vwind10m	2	6.496; 1.928	93.6%
sur_vis	2	6.621; 1.400	89.1%

Tabel 4.3 menunjukkan bahwa jumlah komponen utama yang terbentuk dari seluruh variabel NWP di Stamet Juanda sebanyak 20 komponen. Hasil reduksi yang terbentuk sebanyak 1 hingga 2 komponen.

**Tabel 4.4** *Eigenvalue* dan Keragaman PC Stamet Ngurai Rai

Variabel	Ngurai Rai		
	Jumlah PC	Eigenvalue	Keragaman Kumulatif
hujan_total	1	7.512	83.5%
dew2m	1	8.473	94.2%
wind_gust	1	8.657	96.2%
hi_cloud	1	8.282	92.0%
low_cloud	1	7.326	81.4%
mid_cloud	1	8.429	93.7%
tot_cloud	1	8.393	93.3%
mslp	1	8.958	99.5%
sur_press	1	8.963	99.6%
rhum2m	2	6.974; 1.771	97.2%
shum2m	1	8.481	94.2%
temp2m	2	7.062; 1.877	99.3%
uwind10m	1	8.932	99.3%
vwind10m	1	8.723	96.9%
sur_vis	3	4.389; 2.089; 0.980	82.9%

Jumlah komponen yang terbentuk pada data NWP di Stamet Ngurah Rai berjumlah 19 komponen (Tabel 4.4). Sedangkan untuk Stamet Soetta jumlah komponen yang terbentuk sebanyak 20 komponen (Tabel 4.5).

**Tabel 4.5** *Eigenvalue* dan Keragaman PC Stamet Soetta

Variabel	Soetta		
	Jumlah PC	Eigenvalue	Keragaman Kumulatif
hujan_total	2	6.701; 1.240	88.2%
dew2m	1	8.402	93.4%
wind_gust	1	8.036	89.3%
hi_cloud	1	8.186	91.0%
low_cloud	2	6.158; 1.342	83,3%
mid_cloud	1	8.158	90.6%
tot_cloud	1	8.174	90.8%
Mslp	1	8,936	99,3%
sur_press	1	8,966	99,6%
rh2m	2	7.688; 1.141	98.1%
sh2m	1	8,387	93,2%
temp2m	2	5.206; 3.547	97.3%
uwind10m	1	8,800	97,8%
vwind10m	1	7,820	86.9%
sur_vis	2	6.318; 1.697	89.1%

### 4.3 Penyusunan Model dengan CART

Klasifikasi curah hujan akan menggunakan pendekatan *Classification and Regression Tree* (CART). Adapun variabel respon yang digunakan yaitu berupa data kategorik sehingga pendekatan CART akan menghasilkan suatu pohon klasifikasi (*classification tree*).

#### 4.3.1 Klasifikasi Curah Hujan Stamet Juanda

Berikut tahapan analisis klasifikasi pohon pada Stamet Juanda

##### 4.3.1.1 Pembentukan Pohon Klasifikasi Maksimal

Pembentukan pohon klasifikasi maksimal merupakan tahap pertama dari metode *Classification Trees*. Dalam tahap ini diawali dengan pemilihan *classifier* atau pemilah yakni dengan

memilih variabel prediktor terpenting yang akan digunakan sebagai pemilah awal. Berdasarkan variabel prediktor yang digunakan dalam penelitian ini, proses pemilihan pemilah dilakukan dengan mengacu pada skor kontribusi terbesar yang dihasilkan dari tiap variabel prediktor. Besar kontribusi yang dihasilkan oleh masing-masing variabel prediktor dapat dilihat dalam Tabel 4.6 berikut

**Tabel 4.6** Skor Variabel Prediktor dalam Pembentukan Pohon Klasifikasi Maksimal Stamet Juanda

Variabel	Skor Variabel	
hujan_1t	100.00	
rhun2m1	87.69	
sur_pres	71.14	
shun2m	69.09	
hi_cloud	65.47	
tot_clou	64.22	
dew2m	61.40	

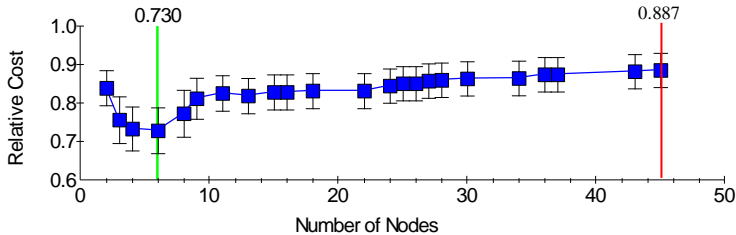
Tabel 4.6 menunjukkan bahwa variabel terpenting yang memiliki kontribusi terbesar adalah variabel hujan\_total1. Selain itu, terdapat beberapa variabel terpenting dan menjadi pemilah utama dalam klasifikasi curah hujan di Stamet Juanda yaitu rhun2m1, sur\_press, shun2m, hi\_cloud, tot\_cloud dan dew2m. Sedangkan variabel lainnya memiliki skor dibawah 50.

Selanjutnya dilakukan proses pemilahan dengan fungsi keheterogenan indeks gini hingga terbentuk pohon klasifikasi maksimal. Pohon klasifikasi maksimal menghasilkan simpul terminal dengan jumlah paling maksimal atau terbanyak dan tingkat kedalaman tertinggi. Jumlah simpul terminal yang dihasilkan oleh pohon klasifikasi maksimal Stamet Juanda adalah sebanyak 45 simpul dengan tingkat kedalaman sebesar 12.

#### 4.3.1.2 Pemangkasan Pohon Klasifikasi Maksimal (Pruning)

Setelah memperoleh pohon klasifikasi maksimal, langkah berikutnya adalah melihat apakah pohon klasifikasi maksimal tersebut perlu dilakukan proses pemangkasan (*pruning*) atau tidak. Proses pemangkasan pohon dilakukan berdasarkan pendekatan *cost complexity minimum* dan setiap hasil pemang-

kasan yang dilakukan memiliki nilai *relative cost* tertentu, kemudian dipilih hasil pemangkasan yang mempunyai nilai *relative cost* paling minimum. Plot *relative cost* dari setiap pemangkasan yang dilakukan pada pohon klasifikasi curah hujan Stamet Juanda sebagai berikut.



**Gambar 4.1** Plot Relative Cost Klasifikasi Curah Hujan Stamet Juanda

Gambar 4.1 menunjukkan bahwa pohon klasifikasi maksimal yang terbentuk dengan jumlah simpul terminal maksimal sebanyak 45 simpul, sehingga perlu dilakukan pemangkasan seperti yang ditunjukkan oleh garis berwarna merah dengan nilai *relative cost* 0.887. Sedangkan garis berwarna hijau dalam Gambar 4.1 menunjukkan bahwa pohon klasifikasi tersebut telah optimal dengan nilai *relative cost* paling minimum yaitu sebesar 0.730. Hasil pemangkasan yang ke-20 tersebut menghasilkan pohon klasifikasi optimal dengan jumlah simpul terminal sebanyak 6 simpul. Urutan mulai dari pohon klasifikasi maksimal yang dibentuk hingga pemangkasan yang dilakukan terhadap pohon tersebut dapat dilihat dalam Tabel 4.7 berikut

**Tabel 4.7** Pembentukan Pohon Klasifikasi Stamet Juanda

<i>Tree Number</i>	<i>Terminal Nodes</i>	<i>Cross-Validated Relative Cost</i>	<i>Resubstitution Relative Cost</i>	<i>Complexity Parameter</i>
1	45	$0.887 \pm 0.045$	0.118	0.000
15	15	$0.830 \pm 0.046$	0.291	0.009
16	13	$0.820 \pm 0.046$	0.318	0.010
17	11	$0.827 \pm 0.046$	0.350	0.012
18	9	$0.813 \pm 0.054$	0.388	0.014
19	8	$0.774 \pm 0.061$	0.410	0.017
20**	6	$0.730 \pm 0.060$	0.485	0.028
21	4	$0.734 \pm 0.057$	0.568	0.031

**Tabel 4.7** Pembentukan Pohon Klasifikasi Stamet Juanda (Lanjutan)

<i>Tree Number</i>	<i>Terminal Nodes</i>	<i>Cross-Validated Relative Cost</i>	<i>Resubstitution Relative Cost</i>	<i>Complexity Parameter</i>
22	3	$0.757 \pm 0.061$	0.625	0.043
23	2	$0.840 \pm 0.046$	0.796	0.129
24	1	$1.000 \pm 0.000$	1.000	0.153






*\*\*pohon klasifikasi optimal*

Tabel 4.7 menunjukkan biaya kesalahan (*cross-validated relative cost*) yang dihasilkan oleh setiap hasil pemangkasan pohon klasifikasi maksimal. Pohon pertama merupakan pembentukan pohon klasifikasi maksimal dengan jumlah simpul terminal sebanyak 45 simpul dan biaya kesalahan sebesar  $0.887 \pm 0.045$ . Sedangkan pohon klasifikasi yang menghasilkan biaya kesalahan paling minimum yaitu sebesar  $0.730 \pm 0.060$  pada saat simpul terminal sebanyak 6.

#### 4.3.1.3 Penentuan Pohon Klasifikasi Optimal

Setelah memperoleh hasil pemangkasan pohon klasifikasi secara iteratif, langkah berikutnya adalah menentukan pohon klasifikasi optimal. Hasil klasifikasi curah hujan di Stamet Juanda diperoleh bahwa pohon klasifikasi optimal dengan jumlah simpul terminal sebanyak 6, dengan nilai *cross validated relative cost* sebesar  $0.730 \pm 0.060$  dan *resubstitution relative cost* sebesar 0.485.

**Tabel 4.8** Skor Variabel Prediktor dalam Pembentukan Pohon Klasifikasi Optimal Stamet Juanda

Variabel	Skor Variabel
hujan_1t	100.00 
rhun2m1	90.16 
shun2m	70.61 
v1wind10	63.62 
dew2m	57.12 

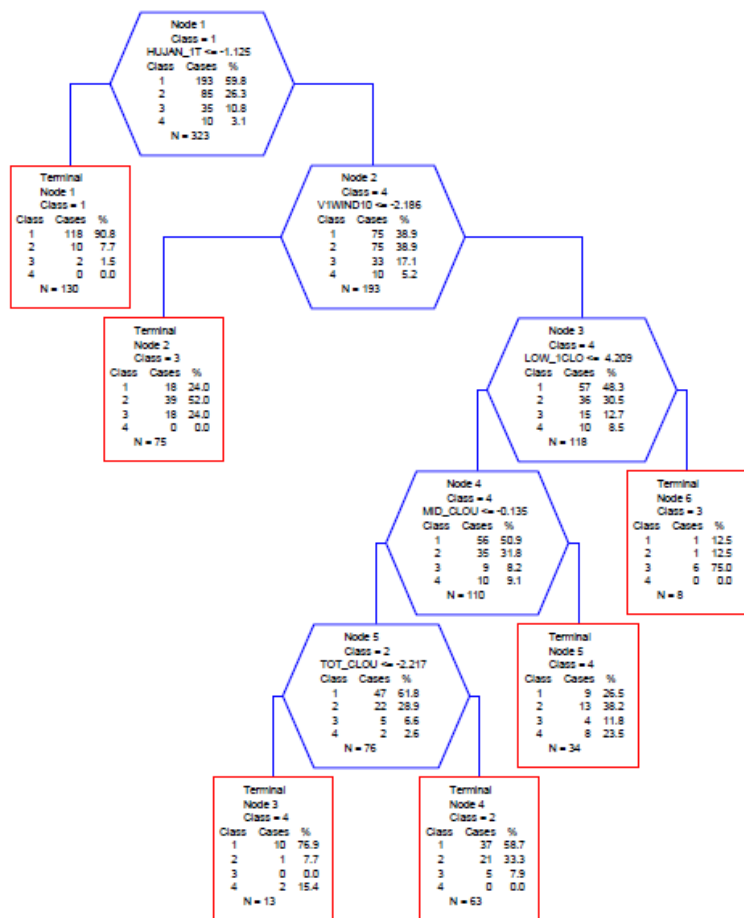
Variabel hujan\_total1 merupakan variabel terpenting dalam pembentukan pohon klasifikasi optimal dengan skor kontribusi 100. Selain itu, terdapat beberapa variabel terpenting dan menjadi pemilah utama dalam klasifikasi curah hujan di Stamet Juanda yaitu rhun2m1, shum2m, vwind10m1 dan dew2m. Sedangkan variabel lainnya memiliki skor dibawah 50 (Tabel 4.8).

Suatu simpul akan terus dipilah menjadi simpul anak baru (kiri dan kanan) sesuai prosedur *binary recursive partitioning*, sampai simpul tersebut telah dianggap memiliki anggota yang homogen atau jika simpul tersebut hanya memiliki 1 anggota pengamatan maka simpul akan menjadi simpul terminal dan tidak akan dipilah lagi. Masing-masing simpul terminal tersebut memiliki karakteristik tertentu sesuai dengan label kelas yang diberikan. Berdasarkan hasil penelusuran 6 simpul terminal tersebut, akan dirangkum pada Tabel 4.9 mengenai pengklasifikasian curah hujan menurut indikasi kesamaan label kelas setiap simpul terminal.

**Tabel 4.9** Kelas Curah Hujan Stamet Juanda pada Masing-masing Simpul Terminal

Kelas	Simpul Terminal Ke-	Persentase (%)
Cerah berawan	1	90.8
Hujan ringan	4	33.3
Hujan sedang	2	24.0
	6	75.0
Hujan lebat	3	15.4
	5	23.5

Tabel 4.9 menunjukkan bahwa dari 6 simpul terminal terbentuk, kelas 1 dan kelas 2 mempunyai masing-masing 1 simpul terminal sedangkan kelas 3 dan kelas 4 mempunyai 2 simpul terminal yang terbentuk. Proses pemilahan mengacu pada pohon klasifikasi optimal yang ditampilkan dalam Gambar 4.2. Pemilahan diawali dengan memilah data sebanyak 323 data berdasarkan variabel hujan\_total1 dengan nilai threshold kurang dari samadengan -1.125. Kemudian data dibagi menjadi dua bagian yaitu simpul kiri (*terminal node* 1) yang terdiri dari 130 data dan simpul kanan (*node* 2) yang terdiri dari 193 data. Pemilahan terus dilakukan secara berulang karena tingkat keheterogenan yang masih tinggi. Pemilahan berhenti dengan jumlah *terminal node* sebanyak 6 simpul.



**Gambar 4.2** Split Plot Pohon Optimal Stamet Juanda

Berikut merupakan karakteristik dari simpul terminal yang akan dijelaskan berdasarkan kategori curah hujan

1. Simpul terminal 1 adalah klasifikasi curah hujan dengan kategori cerah berawan dengan nilai hujan\_total1 kurang dari samadengan -1.125.

2. Simpul terminal 2 adalah klasifikasi curah hujan dengan kategori hujan sedang dengan nilai Vwind10m1 kurang dari samadengan -2.186 dan hujan\_total1 lebih dari -1.125.
3. Simpul terminal 3 adalah klasifikasi curah hujan lebat dengan nilai tot\_cloud kurang dari samadengan -2.217, mid\_cloud kurang dari samadengan -0.135, low\_cloud1 kurang dari samadengan 4.209, vwind10m1 lebih dari -2.186 dan hujan\_total1 lebih dari -1.125.
4. Simpul terminal 4 adalah klasifikasi curah hujan dengan kategori hujan ringan dengan nilai tot\_cloud lebih dari -2.217, mid\_cloud kurang dari samadengan -0.135, low\_cloud1 kurang dari samadengan 4.209, vwind10m1 lebih dari -2.186 dan hujan\_total1 lebih dari -1.125.
5. Simpul terminal 5 adalah klasifikasi curah hujan dengan kategori hujan lebat dengan nilai mid\_cloud lebih dari -0.135, low\_cloud1 kurang dari samadengan 4.209, vwind10m1 lebih dari -2.186 dan hujan\_total1 lebih dari -1.125.
6. Simpul terminal 6 adalah klasifikasi curah hujan dengan kategori hujan ringan dengan nilai low\_cloud1 lebih dari 4.209, vwind10m1 lebih dari -2.186 dan hujan\_total1 lebih dari -1.125.

#### 4.3.1.4 Hasil Ketepatan Klasifikasi Pohon

Tingkat keakuratan hasil klasifikasi pohon yang optimal dihasilkan dari data *learning* dapat dihitung berdasarkan Tabel 4.10

**Tabel 4.10** Ketepatan Klasifikasi Data *Learning* Stamet Juanda

Aktual	Prediksi				Total
	1	2	3	4	
1	118	37	19	19	193
2	10	21	40	14	85
3	2	5	24	4	35
4	0	0	0	10	10
Total	130	63	83	47	323



Berdasarkan informasi pada Tabel 4.10, maka dapat dihitung ketepatan klasifikasi data *learning* sebagai berikut:

$$1 - \text{APER} = \frac{118 + 21 + 24 + 10}{323} \times 100\% = 53.56\%$$

diperoleh ketepatan klasifikasi data *learning* Stamet Juanda sebesar 53.56 %. Namun pohon klasifikasi optimal yang terbentuk perlu divalidasi untuk mengetahui apakah pohon klasifikasi tersebut layak dan dapat digunakan untuk mengklasifikasikan data baru. Tingkat keakuratan hasil klasifikasi pohon yang optimal dihasilkan dari data *testing* dapat dihitung berdasarkan Tabel 4.11

**Tabel 4.11** Ketepatan Klasifikasi Data *Testing* Stamet Juanda

Aktual	Prediksi				Total
	1	2	3	4	
1	8	2	3	0	13
2	5	4	2	0	11
3	1	1	2	0	4
4	2	0	0	0	2
Total	16	7	7	0	30

Ketepatan klasifikasi untuk data *testing* sebagai berikut:

$$1 - \text{APER} = \frac{8 + 4 + 2 + 0}{30} \times 100\% = 46.67\%$$

diperoleh ketepatan klasifikasi untuk data *testing* Stamet Juanda sebesar 46.67%.


### 4.3.2 Klasifikasi Curah Hujan Stamet Ngurah Rai

Tahapan analisis klasifikasi pohon pada Stamet Ngurai Rai yaitu sebagai berikut

#### 4.3.2.1 Pembentukan Pohon Klasifikasi Maksimal

Besar kontribusi yang dihasilkan oleh masing-masing variabel prediktor dapat dilihat dalam Tabel 4.12 berikut

**Tabel 4.12** Skor Variabel Prediktor dalam Pembentukan Pohon Klasifikasi Maksimal Stamet Ngurai Rai

Variabel	Skor Variabel
temp2m1	100.00 

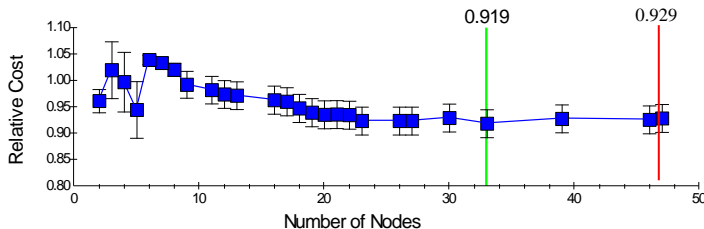
**Tabel 4.12** Skor Variabel Prediktor dalam Pembentukan Pohon Klasifikasi Maksimal Stamet Ngurai Rai (Lanjutan)

Variabel	Skor Variabel	
sur_pres	82.21	
rhumb2m1	76.46	
mslp	70.01	
low_clou	51.68	
hujan_to	50.71	
hi_cloud	50.58	
sur_vis2	50.47	

Tabel 4.12 menunjukkan bahwa variabel terpenting yang memiliki kontribusi terbesar adalah variabel temp2m1. Kemudian sur\_press, rhumb2m1 hingga sur\_vis2 yang menjadi variabel penting dan menjadi pemilah utama lainnya. Sedangkan variabel lainnya memiliki skor dibawah 50. Jumlah simpul terminal yang dihasilkan oleh pohon klasifikasi maksimal Stamet Ngurai Rai adalah sebanyak 47 simpul dengan tingkat kedalaman sebesar 12.

#### 4.3.2.2 Pemangkasan Pohon Klasifikasi Maksimal (Prunning)

Plot *relative cost* dari setiap pemangkasan yang dilakukan pada pohon klasifikasi curah hujan Stamet Ngurai Rai sebagai berikut.



**Gambar 4.3** Plot Relative Cost Klasifikasi Curah Hujan Stamet Ngurai Rai

Gambar 4.3 menunjukkan bahwa pohon klasifikasi maksimal yang terbentuk dengan jumlah simpul terminal maksimal sebanyak 47 simpul, sehingga perlu dilakukan pemangkasan seperti yang ditunjukkan oleh garis berwarna merah dengan nilai *relative cost* 0.929. Sedangkan garis berwarna hijau

dalam Gambar 4.4 menunjukkan bahwa pohon klasifikasi tersebut telah optimal dengan nilai *relative cost* paling minimum yaitu sebesar 0.919. Hasil pemangkasan yang ke-4 tersebut menghasilkan pohon klasifikasi optimal dengan jumlah simpul terminal sebanyak 33 simpul. Urutan mulai dari pohon klasifikasi maksimal yang dibentuk hingga pemangkasan yang dilakukan terhadap pohon tersebut dapat dilihat dalam Tabel 4.13 berikut.

**Tabel 4.13** Pembentukan Pohon Klasifikasi Stamet Ngurai Rai

<i>Tree Number</i>	<i>Terminal Nodes</i>	<i>Cross-Validated Relative Cost</i>	<i>Resubstitution Relative Cost</i>	<i>Complexity Parameter</i>
1	47	$0.929 \pm 0.026$	0.047	0.000
4**	33	$0.919 \pm 0.026$	0.070	0.002
18	11	$0.983 \pm 0.026$	0.249	0.016
19	9	$0.993 \pm 0.025$	0.295	0.018
20	8	$1.022 \pm 0.012$	0.320	0.020
21	7	$1.034 \pm 0.011$	0.350	0.024
22	6	$1.040 \pm 0.009$	0.383	0.026
23	5	$0.945 \pm 0.054$	0.428	0.036
24	4	$0.998 \pm 0.057$	0.514	0.069
25	3	$1.021 \pm 0.054$	0.628	0.091
26	2	$0.962 \pm 0.022$	0.750	0.098
27	1	$1.000 \pm 0.000$	1.000	0.200

\*\*pohon klasifikasi optimal

Tabel 4.13 menunjukkan biaya kesalahan (*cross-validated relative cost*) yang dihasilkan oleh setiap hasil pemangkasan pohon klasifikasi maksimal. Pohon pertama merupakan pembentukan pohon klasifikasi maksimal dengan jumlah simpul terminal sebanyak 47 simpul dan biaya kesalahan sebesar  $0.929 \pm 0.026$ . Sedangkan pohon klasifikasi yang menghasilkan biaya kesalahan paling minimum yaitu sebesar  $0.919 \pm 0.026$  pada saat simpul terminal sebanyak 33.

#### 4.3.2.3 Penentuan Pohon Klasifikasi Optimal

Hasil klasifikasi curah hujan di Stamet Ngurai Rai diperoleh bahwa pohon klasifikasi optimal dengan jumlah simpul terminal sebanyak 33, dengan nilai *cross validated relative cost* sebesar  $0.919 \pm 0.026$  dan *resubstitution relative cost* sebesar 0.070.

**Tabel 4.14** Skor Variabel Prediktor dalam Pembentukan Pohon Klasifikasi Optimal Stamet Ngurai Rai

Variabel	Skor Variabel	
temp2m1	100.00	<div><div></div></div>
sur_pres	81.63	<div><div></div></div>
rhum2m1	75.87	<div><div></div></div>
mslp	69.38	<div><div></div></div>
low_clou	51.10	<div><div></div></div>
hujan_to	50.68	<div><div></div></div>
sur_vis2	50.08	<div><div></div></div>

Variabel temp2m1 merupakan variabel terpenting dalam pembentukan pohon klasifikasi optimal Stamet Ngurah Rai dengan skor kontribusi 100. Selain itu, terdapat beberapa variabel terpenting dan menjadi pemilah utama dalam klasifikasi curah hujan di Stamet Ngurai Rai yang tercantum pada Tabel 4.14. Sedangkan variabel lainnya memiliki skor dibawah 50.

Masing-masing simpul terminal tersebut memiliki karakteristik tertentu sesuai dengan label kelas yang diberikan. Berdasarkan hasil penelusaran 33 simpul terminal pada Lampiran 9, akan dirangkum pada Tabel 4.15 mengenai pengklasifikasian curah hujan menurut indikasi kesamaan label kelas setiap simpul terminal.

**Tabel 4.15** Kelas Curah Hujan Stamet Ngurai Rai pada Masing-masing Simpul Terminal

Kelas	Simpul Terminal Ke-	Persentase (%)	Simpul Terminal Ke-	Persentase (%)
1	1	100	25	88.9
	2	100	27	100
	4	100	29	97.8
	5	100	31	100
	8	100	33	100
	11	100		
2	3	66.7	19	37.5
	6	66.7	20	33.3
	7	84.6	22	100
	9	100	23	78.5
	12	69.2	28	75

**Tabel 4.15** Kelas Curah Hujan Stamet Ngurai Rai pada Masing-masing Simpul Terminal (Lanjutan)

Kelas	Simpul Terminal Ke-	Persentase (%)	Simpul Terminal Ke-	Persentase (%)
3	15	33.3	30	71.4
	16	55.6	32	66.7
	18	100		
	10	100	21	71.4
	13	33.3	26	100
	17	33.3		
4	14	62.5		
5	24	100		

Kelas 2 memiliki jumlah simpul terminal terbanyak dibandingkan dengan kelas lainnya (Tabel 4.15). Proses pemilahan mengacu pada pohon klasifikasi optimal yang ada pada Lampiran 9. Pemilahan diawali dengan memilah data sebanyak 323 data berdasarkan variabel *rhun2m1* dengan nilai *threshold* lebih dari samadengan 1.371. Kemudian data dibagi menjadi dua bagian yaitu simpul kiri (*node* 2) yang terdiri dari 220 data dan simpul kiri (*node* 24) yang terdiri dari 103 data. Pemilahan terus dilakukan secara berulang karena tingkat keheterogenan yang masih tinggi. Pemilahan berhenti dengan jumlah *terminal node* sebanyak 33 simpul.

#### 4.3.2.4 Hasil Ketepatan Klasifikasi Pohon

Tingkat keakuratan hasil klasifikasi pohon yang optimal dihasilkan dari data *learning* dapat dihitung berdasarkan Tabel 4.16

**Tabel 4.16** Ketepatan Klasifikasi Data *Learning* Stamet Ngurai Rai

Aktual	Prediksi					Total
	1	2	3	4	5	
1	174	36	3	2	0	215
2	2	82	5	1	0	90
3	0	0	11	0	0	11
4	0	0	0	5	0	5
5	0	0	0	0	2	2
Total	176	118	19	8	2	323

Berdasarkan informasi pada Tabel 4.16, maka dapat dihitung ketepatan klasifikasi data *learning* Stamet Ngurai Rai sebagai berikut:

$$1 - \text{APER} = \frac{174 + 82 + 11 + 5 + 2}{323} \times 100\% = 84.83 \%$$

diperoleh ketepatan klasifikasi data *learning* sebesar 84.83 %. Sedangkan untuk data *testing* (Tabel 4.17) diperoleh ketepatan klasifikasi sebesar 73.33%.

**Tabel 4.17** Ketepatan Klasifikasi Data *Testing* Stamet Ngurai Rai

Aktual	Prediksi					Total
	1	2	3	4	5	
1	15	2	0	0	0	17
2	4	5	2	0	0	11
3	0	0	2	0	0	2
4	0	0	0	0	0	0
5	0	0	0	0	0	0
Total	19	7	4	1	0	30

$$1 - \text{APER} = \frac{15 + 5 + 2 + 0 + 0}{30} \times 100\% = 73.33\%$$

**4.3.3 Klasifikasi Curah Hujan Stamet Soetta**

Berikut tahapan analisis klasifikasi pohon pada Stamet Soetta.

**4.3.3.1 Pembentukan Pohon Klasifikasi Maksimal**

Besar kontribusi yang dihasilkan oleh masing-masing variabel prediktor disajikan dalam Tabel 4.18 berikut

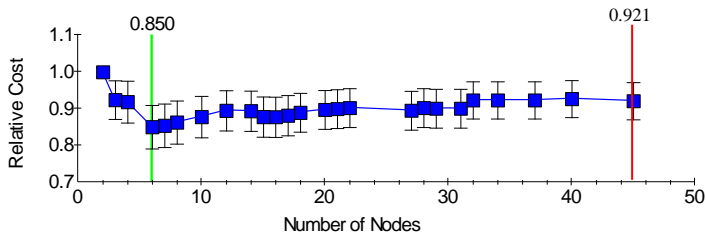
**Tabel 4.18** Skor Variabel Prediktor dalam Pembentukan Pohon Klasifikasi Maksimal Stamet Soetta

Variabel	Skor Variabel
tot_clou	100.00
rhum2m2	90.50
mid_clou	84.94
wind_gus	60.32
hi_cloud	55.14
rhum2m1	51.89

Tabel 4.18 menunjukkan bahwa variabel terpenting yang memiliki kontribusi terbesar adalah variabel *tot\_cloud*. Selain itu, terdapat beberapa variabel terpenting dan menjadi pemilah utama dalam klasifikasi curah hujan di Stamet Soetta yaitu *rum2m2*, *mid\_cloud*, *wind\_gust*, *hi\_cloud* dan *rum2m1*. Sedangkan variabel lainnya memiliki skor dibawah 50. Jumlah simpul terminal yang dihasilkan oleh pohon klasifikasi maksimal Stamet Soetta adalah sebanyak 45 simpul dengan tingkat kedalaman sebesar 13.

#### 4.3.3.2 Pemangkasan Pohon Klasifikasi Maksimal (Pruning)

Plot *relative cost* dari setiap pemangkasan yang dilakukan pada pohon klasifikasi curah hujan Stamet Soetta sebagai berikut.



**Gambar 4.4** Plot Relative Cost Klasifikasi Curah Hujan Stamet Soetta

Gambar 4.8 menunjukkan bahwa pohon klasifikasi maksimal yang terbentuk dengan jumlah simpul terminal maksimal sebanyak 45 simpul, sehingga perlu dilakukan pemangkasan seperti yang ditunjukkan oleh garis berwarna merah dengan nilai *relative cost* 0.921. Sedangkan garis berwarna hijau dalam Gambar 4.8 menunjukkan bahwa pohon klasifikasi tersebut telah optimal dengan nilai *relative cost* paling minimum yaitu sebesar 0.850. Hasil pemangkasan yang ke-22 tersebut menghasilkan pohon klasifikasi optimal dengan jumlah simpul terminal sebanyak 6 simpul. Urutan mulai dari pohon klasifikasi maksimal yang dibentuk hingga pemangkasan yang dilakukan terhadap pohon tersebut dapat dilihat dalam Tabel 4.19 berikut

**Tabel 4.19** Pembentukan Pohon Klasifikasi Stamet Soetta

<i>Tree Number</i>	<i>Terminal Nodes</i>	<i>Cross-Validated Relative Cost</i>	<i>Resubstitution Relative Cost</i>	<i>Complexity Parameter</i>
1	45	$0.921 \pm 0.050$	0.076	0.000
17	14	$0.894 \pm 0.055$	0.212	0.013
18	12	$0.895 \pm 0.055$	0.245	0.013
19	10	$0.878 \pm 0.056$	0.279	0.014
20	8	$0.863 \pm 0.059$	0.326	0.019
21	7	$0.854 \pm 0.059$	0.350	0.020
22**	6	$0.850 \pm 0.059$	0.375	0.020
23	4	$0.918 \pm 0.057$	0.460	0.034
24	3	$0.924 \pm 0.053$	0.583	0.099
25	2	$0.999 \pm 0.012$	0.750	0.133
26	1	$1.000 \pm 0.000$	1.000	0.200

*\*\*pohon klasifikasi optimal*

Tabel 4.19 menunjukkan biaya kesalahan (*cross-validated relative cost*) yang dihasilkan oleh setiap hasil pemangkasan pohon klasifikasi maksimal. Pohon pertama merupakan pembentukan pohon klasifikasi maksimal dengan jumlah simpul terminal sebanyak 45 simpul dan biaya kesalahan sebesar  $0.921 \pm 0.050$ . Sedangkan pohon klasifikasi yang menghasilkan biaya kesalahan paling minimum yaitu sebesar  $0.850 \pm 0.059$  pada saat simpul terminal sebanyak 6.

**4.3.3.3 Penentuan Pohon Klasifikasi Optimal**

Hasil klasifikasi curah hujan di Stamet Soetta diperoleh bahwa pohon klasifikasi optimal dengan jumlah simpul terminal sebanyak 6, dengan nilai *cross validated relative cost* sebesar  $0.850 \pm 0.059$  dan *resubsition relative cost* sebesar 0.375.

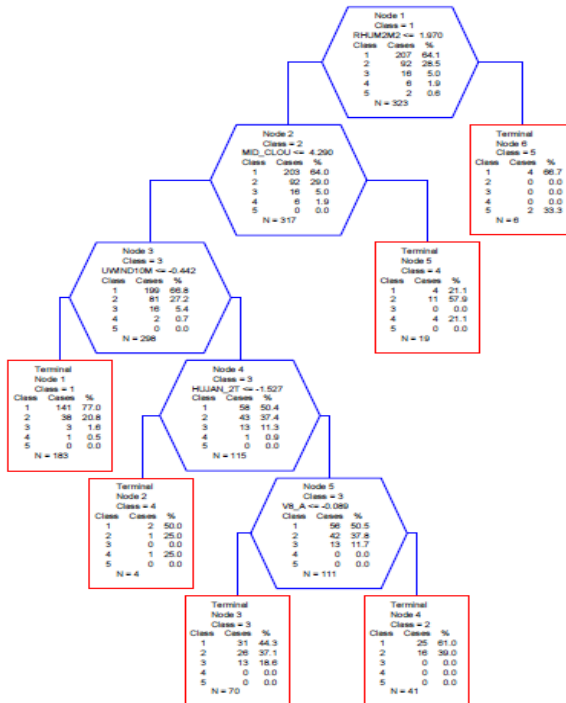
**Tabel 4.20** Skor Variabel Prediktor dalam Pembentukan Pohon Klasifikasi Optimal Stamet Soetta

Variabel	Skor Variabel
mid_clou	100.00
rhum2m2	98.53
tot_clou	80.39

Variabel mid\_cloud merupakan variabel terpenting dalam pembentukan pohon klasifikasi optimal Stamet Soetta dengan skor kontribusi 100.



Pohon klasifikasi yang optimal terdiri dari 6 simpul seperti pada Gambar 4.5. Masing-masing simpul terminal tersebut memiliki karakteristik tertentu sesuai dengan label kelas yang diberikan.



**Gambar 4.5** Split Plot Pohon Optimal Stamet Soetta

Berdasarkan hasil penelusuran 6 simpul terminal tersebut, akan dirangkum pada Tabel 4.21 mengenai pengklasifikasian curah hujan menurut indikasi kesamaan label kelas setiap simpul terminal.

**Tabel 4.21** Kelas Curah Hujan Stamet Soetta pada Masing-masing Simpul Terminal

Kelas	Simpul Terminal Ke-	Persentase (%)
1	1	77.0
2	4	39.0
3	3	18.6
4	2	25.0
	5	21.1
5	6	33.3

Tabel 4.21 menunjukkan bahwa dari 6 simpul terminal terbentuk, kelas 4 mempunyai simpul yang paling banyak terbentuk. Proses pemilahan mengacu pada pohon klasifikasi optimal yang ditampilkan dalam Gambar 4.12. Pemilahan diawali dengan memilah data sebanyak 323 data berdasarkan variabel *rhun2m2* dengan nilai threshold  $\leq 1.970$  Kemudian data dibagi menjadi dua bagian yaitu simpul kiri (*node 2*) yang terdiri dari 317 data dan simpul kiri (*terminal node 6*) yang terdiri dari 6 data. Pemilahan terus dilakukan secara berulang karena tingkat keheterogenan yang masih tinggi. Pemilahan berhenti dengan jumlah *terminal node* sebanyak 6 simpul. Berikut merupakan karakteristik dari 6 simpul terminal pohon klasifikasi optimal Stamet Soetta

1. Simpul terminal 1 adalah klasifikasi curah hujan dengan kategori cerah berawan dengan nilai *uwind10m* kurang dari samadengan -0.442; *mid\_cloud* kurang dari samadengan 4.290 dan *rhun2m2* kurang dari samadengan 1.970.
2. Simpul terminal 2 adalah klasifikasi curah hujan dengan kategori hujan lebat dengan nilai hujan\_tot2 kurang dari samadengan =1.527; *uwind10m* lebih dari -0.442; *mid\_cloud* kurang dari samadengan 4.290 dan *rhun2m2* kurang dari samadengan 1.970.
3. Simpul terminal 3 adalah klasifikasi curah hujan dengan kategori hujan sedang dengan nilai *sur\_vis3* kurang dari samadengan -0.089; hujan\_tot2 lebih dari -1.527; *uwind10m*

- lebih dari -0.442; mid\_cloud kurang dari samadengan 4.290 dan rhum2m2 kurang dari samadengan 1.970.
4. Simpul terminal 4 adalah klasifikasi curah hujan dengan ketagori hujan ringan dengan nilai sur\_vis3 lebih dari -0.089; hujan\_tot2 lebih dari -1.527; uwind10m lebih dari -0.442; mid\_cloud kurang dari samadengan 4.290 dan rhum2m2 kurang dari samadengan 1.970.
  5. Simpul terminal 5 adalah klasifikasi curah hujan dengan ketagori hujan lebat dengan nilai mid\_cloud lebih dari 4.290 dan rhum2m2 kurang dari samadengan 1.970.
  6. Simpul terminal 6 adalah klasifikasi curah hujan dengan kategori hujan lebat sekali dengan nilai rhum2m2 lebih dari 1.970.

#### 4.3.3.4 Hasil Ketepatan Klasifikasi Pohon

Tingkat keakuratan hasil klasifikasi pohon yang optimal dihasilkan dari data *learning* dapat dihitung berdasarkan Tabel 4.22

**Tabel 4.22** Ketepatan Klasifikasi Data *Learning* Stamet Soetta

Aktual	Prediksi					Total
	1	2	3	4	5	
1	141	25	31	6	4	207
2	38	16	26	12	0	92
3	3	0	13	0	0	16
4	1	0	0	5	0	6
5	0	0	0	0	2	2
Total	183	41	70	23	6	323

dapat dihitung ketepatan klasifikasi data *learning* sebagai berikut:

$$1 - \text{APER} = \frac{141 + 16 + 13 + 5 + 2}{323} \times 100\% = 54.79\%$$

diperoleh ketepatan klasifikasi data *learning* Stamet Soetta sebesar 54.79% dan untuk data *testing* dapat dihitung berdasarkan Tabel 4.23

**Tabel 4.23** Ketepatan Klasifikasi Data *Testing* Stamet Soetta

Aktual	Prediksi					Total
	1	2	3	4	5	
1	14	6	0	0	0	20
2	2	6	0	1	0	9
3	0	1	0	0	0	1
4	0	0	0	0	0	0
5	0	0	0	0	0	0
Total	16	13	0	1	0	0

Ketepatan klasifikasi untuk data *testing* sebagai berikut:

$$1 - \text{APER} = \frac{14 + 6 + 0 + 0 + 0}{30} \times 100\% = 66.67 \%$$

diperoleh ketepatan klasifikasi untuk data *testing* Stamet Soetta sebesar 66.67%.

#### 4.4 Penyusunan Model *Random Forests*

Analisis *random forests* dilakukan pembentukan pohon klasifikasi dengan jumlah pohon yang dibentuk dicobakan pada kombinasi sebanyak 5, 10, 15, 20 dan 25 pohon. Pohon yang dihasilkan merupakan pohon klasifikasi optimal karena tanpa melewati proses pemangkasan (*prunning*). Berikut hasil akurasi klasifikasi curah hujan dari masing-masing kombinasi yang dicobakan untuk Stamet Juanda

**Tabel 4.24** Hasil Klasifikasi *Random Forests* Stamet Juanda

Jumlah Pohon	Data	1-APER (%)
5	Data Learning	96.59
	Data Testing	43.33
10	Data Learning	98.76
	Data Testing	43.33
15	Data Learning	99.07
	Data Testing	50.00
20*	Data Learning	99.38
	Data Testing	53.33
25	Data Learning	100.00
	Data Testing	53.33

\*kombinasi jumlah pohon terpilih

Tabel 4.24 menunjukkan bahwa kombinasi jumlah pohon yang optimal pada saat pohon dicobakan sebanyak 20 pohon. Hal ini berarti hasil ketepatan klasifikasi menggunakan *random forests* untuk Stamet Juanda pada data *learning* sebesar 99.38 persen dan data *testing* sebesar 53.33 persen. Sedangkan pada Stamet Ngurai Rai disajikan dalam Tabel 4.25

**Tabel 4.25** Hasil Klasifikasi *Random Forests* Stamet Ngurai Rai

Jumlah Pohon	Data	1-APER (%)
5	Data Learning	96.28
	Data Testing	43.33
10	Data Learning	99.38
	Data Testing	56.67
15	Data Learning	99.07
	Data Testing	53.33
20*	Data Learning	99.38
	Data Testing	53.33
25	Data Learning	100.0
	Data Testing	53.33

\*kombinasi jumlah pohon terpilih

Kombinasi jumlah pohon yang optimal pada saat pohon dicobakan untuk Stamet Ngurai Rai sebanyak 20 pohon dengan hasil ketepatan klasifikasi data *learning* sebesar 99.38 persen dan data *testing* sebesar 53.33 persen. Hasil akurasi klasifikasi curah hujan pada Stamet Soetta sebagai berikut

**Tabel 4.26** Hasil Klasifikasi *Random Forests* Stamet Soetta

Jumlah Pohon	Data	1-APER (%)
5	Data Learning	97.21
	Data Testing	56.67
10	Data Learning	98.45
	Data Testing	56.67
15	Data Learning	99.69
	Data Testing	53.33
20*	Data Learning	99.69
	Data Testing	70.00
25	Data Learning	99.69
	Data Testing	70.00

\*kombinasi jumlah pohon terpilih

Tabel 4.26 menunjukkan bahwa kombinasi jumlah pohon yang optimal pada saat pohon dicobakan sebanyak 20 pohon. Hasil ketepatan klasifikasi untuk data *learning* sebesar 99.69 persen dan data *testing* sebesar 70.00 persen.

#### 4.5 Kinerja Model

Kriteria yang digunakan untuk membandingkan ketepatan klasifikasi CART dan *random forests* adalah *total accuracy rate* (1-APER). Metode klasifikasi yang memiliki kriteria ketepatan klasifikasi yang lebih baik dipilih sebagai metode yang terbaik untuk meng-klasifikasikan curah hujan. Berikut perbandingan hasil klasifikasi antara CART dan *random forests*.

**Tabel 4.27** Perbandingan Hasil Klasifikasi antara CART dan *Random Forests*

Stasiun Meteorologi	Metode Klasifikasi	1-APER(%)
Juanda	CART	Data Learning
		53.56
	Random Forests	Data Testing
		46.67
Ngurai Rai	CART	Data Learning
		99.38
	Random Forests	Data Testing
		53.33
Soetta	CART	Data Learning
		84.83
	Random Forests	Data Testing
		73.33
Soetta	CART	Data Learning
		54.79
	Random Forests	Data Testing
		66.67
Soetta	CART	Data Learning
		99.69
	Random Forests	Data Testing
		70.00

Tabel 4.27 menunjukkan bahwa pada Stamet Juanda nilai akurasi data *learning* metode CART dan *random forests* memiliki nilai yang lebih tinggi dari akurasi data *testing*. Hal ini berarti ada kasus *underfitting* pada data Stamet Juanda. Begitu juga dengan Stamet Ngurai Rai, ada kasus *underfitting* pada nilai akurasi yang dihasilkan untuk data *learning*. Untuk Stamet Soetta, pada metode CART terdapat kasus *overfitting* dimana nilai akurasi data *testing* lebih tinggi dari data *learning*. Hasil akurasi dari kedua

metode ini, diperoleh bahwa akurasi data *testing* dengan metode *random forests* memiliki nilai yang lebih tinggi dari akurasi metode CART. Oleh karena itu, dapat dikatakan bahwa metode *random forests* mampu meningkatkan hasil akurasi dalam prakiraan curah hujan harian.

*(halaman ini sengaja dikosongkan)*



## **BAB V**

### **KESIMPULAN DAN SARAN**

#### **5.1 Kesimpulan**

Berdasarkan analisis yang dilakukan, diperoleh kesimpulan sebagai berikut:

1. Hasil reduksi dimensi menggunakan metode PCA menghasilkan total komponen utama (PC) pada Stasiun Meteorologi sebanyak 20 komponen, Stasiun Meteorologi Ngurai Rai sebanyak 19 komponen dan pada Stasiun Meteorologi Soekarno-Hatta (Soetta) sebanyak 20 komponen.
2. Hasil ketepatan klasifikasi untuk metode CART pada Stasiun Meteorologi Juanda, Stasiun Meteorologi Ngurai Rai dan pada Stasiun Meteorologi Soekarno-Hatta (Soetta) yaitu 46.67%, 73.33% dan 66.67%. Sedangkan untuk metode *random forests* ketepatan klasifikasi pada Stasiun Meteorologi Juanda, Stasiun Meteorologi Ngurai Rai dan pada Stasiun Meteorologi Soekarno-Hatta (Soetta) yaitu 53.33%, 53.33% dan 70.00%.
3. *Random forests* lebih mampu meningkatkan hasil akurasi prakiraan curah hujan.

#### **5.2 Saran**

Penelitian selanjutnya disarankan untuk mengatasi data yang *missing* pada variabel curah hujan dan NWP sehingga hasil klasifikasi yang diperoleh akan semakin baik dan dapat menggunakan periode data *learning* dan *testing* yang lebih panjang. Selain itu, metode *random forests* dapat terus digunakan untuk mengklasifikasikan curah hujan dengan menggunakan kombinasi pohon yang lebih besar.

*(halaman ini sengaja dikosongkan)*

## DAFTAR PUSTAKA

- Arfianto. (2006). *Aplikasi Model Regresi Logistik untuk Prakiraan Kejadian Hujan*. Bogor: IPB.
- Ayuni, N. W. (2013). *Reduksi Dimensi Luaran NWP dengan Menggunakan Kernel Sliced Inverse Regression (KSIR) sebagai Pra-Processing pada Model Output Statistics*. Surabaya: ITS.
- Azmi, U. (2017). *Prediksi Curah Hujan melalui Model Output Statistics menggunakan Classification and Regression Trees dengan Pre-Processing Principal Component Analysis*. Surabaya: Tugas Akhir Jurusan Statistika ITS.
- Breiman, L. (2001). *Random Forests*. Berkeley: Statistics Department University of California.
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1993). *Classification and Regression Trees*. USA: Chapman & Hall.
- Bremen, L. V. (2007). Combination of Deterministic and Probabilistic Meteorological Models to Enhance Wind Farm Power Forecast. *Journal of Physic*.
- Dewi, N. K., Syafitri, U. D., & Mulyadi, S. Y. (2011). Penerapan Metode Random Forests dalam Driver Analysis. *Forum Statistika dan Komputasi*.
- Glahn, H. R., & Lowry, D. A. (1972). The Use of Model Output Statistics in Objective Weather Forecasting. *National Weather Service*, 1203-1211.
- Hutapea, T. D., Kurniawan, R., & Hanggoro, W. (2016). Pengaruh Topografi dan Luas Daratan Model WRF terhadap Hasil Prediksi Temperatur Permukaan Di Wilayah Kepulauan Indonesia. *Jurnal Meteorologi dan Geofisika*, 16 No. 3, 179-187.
- Johnson, R. A., & Wichern, D. W. (2007). *Applied Multivariate statistical Analysis Sixth Edition*. New Jersey: Prentice Hall.

- Kusumawardani, R. (2014). *Prediksi Suhu Maksimum, Suhu Minimum dan Kelembapan Rata-rata Relatif dalam Jangka Pendek dengan Multivariate Regression melalui Pra-Pemrosesan PCA*. Surabaya: Jurusan Statistika ITS.
- Lewis, R. J. (2000). *An Introduction to Classification and Regression Tree Analysis, Annual Meeting of The Society for Academic Emergency Medicine*. San Fransisco.
- Nichols, M. (2008). Model Output Statistics. *Independent Research Program*.
- Priambudi, A. K. (2006). *Regresi Splines Adaptif Berganda untuk Peramalan Suhu dan Kelembapan*. Surabaya: Tugas Akhir Jurusan Statistika ITS.
- Safitri, R. (2012). Model Output Statistics dengan Projection Pursuit Regression untuk Meramalkan Suhu Minimu, Suhu Maksimum dan Kelembapan. *Jurnal Sains dan Seni ITS Vol 1 Nomer 1*.
- Sari, G. A. (2013). *Pendekatan Regresi Partial Least Squares Univariate dan Multivariate Response untuk Memprediksi suhu dan Kelembapan*. Surabaya: ITS.
- Setiarnoputera, M. L. (2015). *Prakiraan Suhu dan Kelembapan di DKI Jakarta menggunakan M-PLS*. Surabaya: Jurusan Statistika ITS.
- Sungkono, J. (2013). Resampling Bootstrap pada R. *Magistra No. 84 Th. XXV*.
- Vaxjo, K. (2014). Evaluation of Logistic Regression and Random Forests Classification Based on Prediction Accuracy and Metadata Analysis. *Institutionen for Matematics*.
- Wardani, I. K. (2010). *Manfaat Prediksi Cuaca Jangka Pendek Berdasarkan Data Ra-diosonde dan Numerical Weather Prediction (NWP) untuk Pertanian Daerah Papua*. Jombang: Universitas Pesantren Tinggi Darul Ulum.
- Wilks, D. S. (2006). *Statistical Methods in The Atmospheric Sciences* (2nd ed.). United States of America: Academic Press.

## LAMPIRAN

### Lampiran 1. Data Observasi Curah Hujan

Waktu	Stasiun Meteorologi		
	Juanda	Ngurai Rai	Soetta
02/01/2015 0:00	43,7	11,9	10,5
03/01/2015 0:00	1,8	6	9,7
04/01/2015 0:00	0	4,4	9999
05/01/2015 0:00	61,2	2,2	5,9
06/01/2015 0:00	5,8	9999	0
07/01/2015 0:00	0	8888	0
:	:	:	:
:	:	:	:
:	:	:	:
16/05/2015 0:00	0	0	11
17/05/2015 0:00	0	0	0
20/05/2015 0:00	0	0	0
21/05/2015 0:00	0	0	0
:	:	:	:
:	:	:	:
:	:	:	:
:	:	:	:
23/11/2016 0:00	8888	8888	1,2
10/12/2016 0:00	0,6	15,1	0

Keterangan :

9999 = Data tidak tercatat (*missing value*)

8888 = Tidak terukur (TTU)

## Lampiran 2. Data NWP Stamet Juanda

Waktu	Grid								
	1_1	0_1	-1_1	1_0	0_0	-1_0	1_-1	0_-1	-1_-1
	===== hujan total =====								
2015:1:2:0	4,55	2,01	2,78	0	5,61	9,61	4,51	5,41	11,86
2015:1:3:0	8,05	8,26	9,19	2,8	8,45	9,6	3,1	2,65	5,68
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	30,05	12,4	11,35	10,76	9,18	6,56	10,99	20,45	22,99
2016:12:10:0	29,24	16,57	15,49	2,89	6,86	12,05	0	3,04	9,91
	===== dew point 2m =====								
2015:1:2:0	24,45	23,24	23,10	24,11	22,62	22,66	21,35	21,07	21,35
2015:1:3:0	24,41	23,26	23,08	24,26	22,71	22,68	21,57	21,14	21,50
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	24,17	23,06	22,55	23,96	22,52	22,03	21,84	21,25	21,10
2016:12:10:0	24,11	23,02	23,02	23,87	22,23	22,06	20,80	20,10	20,32
	===== wind gust =====								

2015:1:2:0	8,88	8,04	7,76	8,18	7,24	7,19	9,26	9,24	7,43
2015:1:3:0	8,28	7,33	6,78	7,30	6,46	6,19	8,07	8,82	7,14
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	2,88	2,84	2,70	3,12	2,70	3,25	3,98	3,30	3,90
2016:12:10:0	5,34	4,95	4,75	5,10	4,53	6,05	4,85	6,99	7,19
	===== high cloud =====								
2015:1:2:0	35,04	31,08	25,68	32,52	27,67	20,27	21,21	17,07	22,21
2015:1:3:0	17,70	17,57	18,78	14,15	12,18	10,78	9,78	9,74	11,13
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	7,48	10,53	16,26	5,18	6,13	7,43	8,44	8,99	6,39
2016:12:10:0	46,29	47,10	44,77	46,75	43,73	42,08	39,58	43,23	46,09
	===== low cloud =====								
2015:1:2:0	3,08	3,47	1,60	8,03	6,33	4,93	6,03	7,15	6,40
2015:1:3:0	0,61	2,54	3,33	0,66	2,66	2,68	0,00	0,00	1,45
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:

2016:11:23:0	0,00	0,12	3,53	0,00	0,00	11,23	0,00	0,36	20,56
2016:12:10:0	5,66	11,22	12,74	5,78	7,70	6,53	5,40	2,29	0,00
	===== middle cloud =====								
2015:1:2:0	5,17	3,42	2,92	0,50	0,42	0,29	1,50	1,17	1,00
2015:1:3:0	5,50	6,33	8,58	1,83	4,29	6,04	1,00	6,67	11,08
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	11,63	10,21	10,50	2,92	1,71	5,96	1,21	1,04	17,21
2016:12:10:0	32,04	25,96	21,17	29,42	19,75	16,83	30,92	21,42	21,29
	===== total cloud =====								
2015:1:2:0	14,65	12,85	10,23	13,91	11,64	8,60	9,72	8,57	10,02
2015:1:3:0	8,05	8,90	10,35	5,64	6,44	6,58	3,68	5,57	7,99
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	6,43	7,01	10,19	2,74	2,63	8,19	3,28	3,52	14,60
2016:12:10:0	28,28	28,36	26,50	27,62	23,98	22,12	25,48	22,61	22,81
	===== mean sea level pressure =====								
2015:1:2:0	1006,83	1006,91	1006,97	1006,77	1006,89	1006,98	1006,72	1006,76	1007,42



2015:1:3:0	1007,02	1007,04	1007,11	1006,95	1007,06	1007,11	1006,94	1006,89	1007,68
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	1007,01	1007,02	1007,00	1007,01	1006,97	1006,98	1007,11	1007,05	1007,20
2016:12:10:0	1006,93	1006,98	1007,02	1006,90	1006,93	1006,98	1006,94	1006,91	1007,53
	===== surface pressure =====								
2015:1:2:0	1005,63	1006,79	1006,60	1007,06	1001,66	995,91	976,25	956,61	952,02
2015:1:3:0	1006,11	1007,20	1006,96	1007,49	1002,05	996,26	976,65	957,05	952,37
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	1005,31	1006,41	1006,11	1006,85	1001,33	995,39	976,31	956,47	951,50
2016:12:10:0	1005,55	1006,67	1006,49	1006,98	1001,54	995,78	976,31	956,68	952,08
	===== relative humidity 2m =====								
2015:1:2:0	82,09	79,28	78,06	77,60	75,34	77,83	72,94	79,17	84,73
2015:1:3:0	83,98	81,77	81,29	80,18	77,40	80,22	73,54	78,00	83,37
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	72,95	70,64	68,76	71,68	70,07	69,89	74,80	79,21	81,65

2016:12:10:0	78,56	78,45	78,37	75,06	72,91	71,59	71,10	71,17	74,16
	===== specific humidity 2m =====								
2015:1:2:0	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02
2015:1:3:0	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02
2016:12:10:0	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02
	===== temperature 2m =====								
2015:1:2:0	29,54	27,95	28,11	29,54	28,00	27,91	27,61	26,42	25,63
2015:1:3:0	29,34	28,03	27,86	29,34	27,84	27,56	28,34	26,96	26,34
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	31,44	33,55	34,15	31,44	31,86	32,10	29,63	28,12	26,36
2016:12:10:0	30,64	28,86	28,81	30,64	28,42	29,01	27,12	27,36	27,15
	===== u wind 10m =====								
2015:1:2:0	7,49	5,74	5,19	6,46	4,20	4,74	5,21	5,65	4,97
2015:1:3:0	6,84	5,11	4,48	5,99	3,76	4,06	4,21	5,08	4,43

:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	-0,48	-0,22	0,35	-1,36	-0,80	-0,29	-1,21	-1,08	-1,02
2016:12:10:0	4,00	2,65	2,41	2,47	1,42	2,08	1,13	2,08	2,31
	===== V wind 10m =====								
2015:1:2:0	-1,65	-1,15	-0,83	-1,38	-0,17	0,19	0,62	2,04	1,78
2015:1:3:0	-2,58	-1,66	-1,13	-2,55	-0,41	-0,04	0,25	2,61	2,56
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	-0,15	-0,92	-0,32	0,23	-0,19	0,89	0,16	0,22	1,92
2016:12:10:0	-0,52	-0,66	0,03	-1,11	0,51	2,34	0,37	3,26	4,88
	===== surface vis =====								
2015:1:2:0	24160,83	24160,83	24160,83	24160,83	24160,83	24160,83	24160,83	24160,83	24160,83
2015:1:3:0	24135,36	24135,36	24135,36	24135,36	24135,36	24135,36	24135,36	24135,36	24135,36
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	24107,61	24107,61	24107,61	24107,61	24107,61	24107,61	23232,61	21374,28	20749,28
2016:12:10:0	21166,28	21824,61	22099,61	23824,61	24207,94	24207,94	24207,94	24207,94	24207,94

### Lampiran 3. Data NWP Stamet Ngurai Rai

Waktu	Grid								
	_1_1	_0_1	_-1_1	_1_0	_0_0	_-1_0	_1_-1	_0_-1	_-1_-1
	===== hujan total =====								
2015:1:2:0	8,31	6,94	11,43	7,84	9,66	11,30	19,55	21,74	19,22
2015:1:3:0	5,91	4,58	8,07	13,90	12,76	10,61	14,66	12,72	10,58
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	6,97	10,45	6,66	0,24	0,34	0,55	0,35	0,42	1,08
2016:12:10:0	1,25	2,37	39,97	48,54	40,65	48,06	56,75	47,89	43,38
	===== dew point 2m =====								
2015:1:2:0	21,93	22,25	23,78	24,53	24,59	24,67	24,81	24,84	24,83
2015:1:3:0	22,22	22,55	24,11	24,75	24,76	24,81	24,91	24,86	24,83
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	22,36	22,35	23,48	23,85	23,87	23,95	24,15	24,14	24,14
2016:12:10:0	22,51	22,64	24,00	24,41	24,40	24,52	24,61	24,56	24,51
	===== wind gust =====								

2015:1:2:0	8,07	7,48	8,17	9,70	9,36	9,31	10,01	9,86	9,75
2015:1:3:0	8,01	7,40	8,31	9,24	8,95	8,86	8,96	8,84	8,80
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	2,50	2,84	3,39	2,58	2,68	2,87	2,47	2,53	2,70
2016:12:10:0	4,01	3,75	5,20	4,97	4,75	5,12	5,80	5,75	5,64
	===== high cloud =====								
2015:1:2:0	32,57	29,65	24,90	28,03	24,01	21,24	33,46	29,76	30,63
2015:1:3:0	20,53	19,36	18,48	18,86	18,98	18,19	20,95	19,87	19,03
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	8,79	5,20	4,73	2,25	3,31	4,09	4,44	5,89	6,33
2016:12:10:0	35,20	34,25	34,76	35,36	35,13	35,10	36,03	37,66	37,18
	===== low cloud =====								
2015:1:2:0	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
2015:1:3:0	0,00	0,29	0,55	1,54	1,36	1,25	3,17	1,46	0,70
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:

2016:11:23:0	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
2016:12:10:0	0,00	1,88	4,01	0,00	1,19	2,36	0,00	0,70	1,98
	===== middle cloud =====								
2015:1:2:0	14,46	10,29	6,42	12,04	13,50	11,92	23,33	23,50	20,88
2015:1:3:0	11,04	10,17	5,92	17,13	16,83	14,00	19,21	18,54	17,42
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	1,38	0,88	1,83	0,79	1,25	2,88	2,83	3,21	4,46
2016:12:10:0	17,17	16,04	11,63	19,96	19,33	18,88	21,08	23,38	24,83
	===== total cloud =====								
2015:1:2:0	15,91	13,53	10,63	13,55	12,67	11,20	19,17	17,99	17,40
2015:1:3:0	10,65	10,08	8,45	12,64	12,52	11,23	14,57	13,42	12,48
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	3,45	2,06	2,21	1,01	1,53	2,36	2,46	3,08	3,63
2016:12:10:0	17,75	17,61	17,02	18,68	18,77	19,00	19,34	20,83	21,62
	===== mean sea level pressure =====								
2015:1:2:0	1005,99	1006,17	1006,00	1005,87	1005,91	1005,91	1005,82	1005,82	1005,86

2015:1:3:0	1006,49	1006,68	1006,30	1006,27	1006,25	1006,15	1006,34	1006,18	1006,17
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	1007,31	1007,33	1007,25	1007,30	1007,29	1007,27	1007,30	1007,28	1007,28
2016:12:10:0	1006,75	1006,88	1006,82	1006,63	1006,75	1006,79	1006,48	1006,63	1006,73
	===== surface pressure =====								
2015:1:2:0	973,25	974,06	988,09	1004,82	1005,28	1006,05	1005,89	1005,90	1006,03
2015:1:3:0	973,94	974,77	988,82	1005,49	1006,00	1006,82	1006,53	1006,61	1006,80
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	974,36	974,95	988,94	1006,15	1006,54	1007,22	1007,33	1007,26	1007,32
2016:12:10:0	973,78	974,45	988,41	1005,32	1005,76	1006,51	1006,39	1006,40	1006,55
	===== relative humidity 2m =====								
2015:1:2:0	77,40	79,20	79,61	77,47	77,72	78,01	78,95	79,01	78,99
2015:1:3:0	78,52	80,91	81,82	78,47	78,50	78,60	78,40	78,30	78,40
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	81,29	81,99	76,95	73,40	73,68	73,75	74,19	74,39	74,44

2016:12:10:0	83,27	83,99	83,74	79,60	79,38	80,05	79,57	79,36	79,67
	===== specific humidity 2m =====								
2015:1:2:0	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02
2015:1:3:0	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02
2016:12:10:0	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02
	===== temperature 2m =====								
2015:1:2:0	27,47	27,33	29,33	29,96	30,01	30,02	29,96	30,01	29,97
2015:1:3:0	27,56	27,22	29,19	29,84	29,84	29,83	29,84	29,84	29,80
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	27,75	27,50	30,52	30,42	30,37	30,34	30,42	30,37	30,34
2016:12:10:0	27,80	27,89	29,60	29,46	29,51	29,54	29,46	29,51	29,54
	===== u wind 10m =====								
2015:1:2:0	4,63	4,73	6,80	8,36	8,18	8,02	8,59	8,46	8,38
2015:1:3:0	4,91	4,88	7,06	8,26	8,05	7,78	7,94	7,91	7,78



:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	-0,47	-0,46	-0,86	-0,64	-0,80	-0,88	-0,79	-0,79	-0,80
2016:12:10:0	1,72	1,57	2,46	1,32	1,58	2,09	-0,08	0,48	1,15
	===== V wind 10m =====								
2015:1:2:0	0,65	0,38	1,03	1,14	0,71	0,95	1,11	1,00	1,09
2015:1:3:0	1,72	1,56	2,60	1,87	1,75	2,24	1,98	1,91	2,01
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	1,75	1,82	3,38	2,28	2,29	2,62	2,02	2,04	2,21
2016:12:10:0	2,18	2,24	4,22	2,79	3,09	4,07	3,13	3,54	4,05
	===== surface vis =====								
2015:1:2:0	24160,83	24160,83	24160,83	24160,83	24160,83	24160,83	24160,83	24160,83	24160,83
2015:1:3:0	24135,36	24135,36	24135,36	24135,36	24135,36	24135,36	24135,36	24135,36	24135,36
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	24107,61	24107,61	24107,61	24107,61	24107,61	24107,61	24107,61	24107,61	24107,61
2016:12:10:0	24207,94	24207,94	23957,94	24207,94	24207,94	24207,94	24207,94	24207,94	24207,94

#### Lampiran 4. Data NWP Stamet Soetta

Waktu	Grid								
	_1_1	_0_1	_-1_1	_1_0	_0_0	_-1_0	_1_-1	_0_-1	_-1_-1
	===== hujan total =====								
2015:1:2:0	9,30	6,54	5,74	0,88	5,23	4,73	3,52	5,79	5,56
2015:1:3:0	6,43	6,51	10,71	1,02	0,68	1,99	0,33	1,38	3,41
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	10,38	7,11	4,71	3,93	0,00	9,66	0,00	5,85	9,63
2016:12:10:0	2,94	2,24	2,14	0,26	0,00	0,00	0,00	0,00	0,00
	===== dew point 2m =====								
2015:1:2:0	24,29	24,23	24,20	23,87	22,34	22,29	21,64	21,82	21,94
2015:1:3:0	24,07	24,00	24,00	23,63	21,97	22,00	21,46	21,54	21,69
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	24,28	24,30	24,35	24,04	22,98	22,99	22,11	22,34	22,35
2016:12:10:0	24,30	24,18	24,14	24,06	22,25	22,07	21,69	21,67	21,73
	===== wind gust =====								

2015:1:2:0	8,59	8,65	8,60	8,67	8,52	8,74	8,85	8,87	9,06
2015:1:3:0	9,60	9,65	10,00	9,77	9,43	9,72	9,65	9,72	9,62
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	3,18	3,39	3,78	3,65	3,38	3,64	3,42	3,58	3,94
2016:12:10:0	7,42	7,13	7,04	7,39	6,82	7,19	7,07	7,03	7,23
	===== high cloud =====								
2015:1:2:0	13,59	14,23	13,66	8,30	10,14	8,88	7,89	5,87	5,67
2015:1:3:0	24,31	25,94	30,34	20,35	22,41	29,08	13,63	21,36	27,40
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	31,57	29,77	27,48	27,90	25,72	25,75	16,35	14,98	15,43
2016:12:10:0	39,85	33,58	39,86	32,35	18,90	25,35	23,56	11,68	11,28
	===== low cloud =====								
2015:1:2:0	0,00	0,00	0,00	0,00	0,00	0,00	1,37	1,32	0,22
2015:1:3:0	0,00	0,00	0,00	0,00	0,00	0,00	0,69	0,00	0,00
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:

2016:11:23:0	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,20	0,43
2016:12:10:0	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	===== middle cloud =====								
2015:1:2:0	11,17	9,17	8,71	12,29	10,63	8,79	13,96	11,21	9,08
2015:1:3:0	24,46	29,25	33,58	21,00	22,63	27,50	24,71	21,13	20,58
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	21,21	17,04	16,71	17,00	16,96	15,83	9,33	13,75	12,58
2016:12:10:0	8,79	8,08	6,83	3,42	3,33	2,67	2,13	1,29	1,42
	===== total cloud =====								
2015:1:2:0	8,35	7,90	7,58	6,93	7,00	5,92	7,78	6,18	5,03
2015:1:3:0	16,45	18,60	21,58	13,94	15,18	19,12	13,10	14,33	16,20
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	17,82	15,83	14,96	15,18	14,45	14,03	8,67	9,73	9,61
2016:12:10:0	16,49	14,14	15,87	12,18	7,56	9,53	8,74	4,40	4,30
	===== mean sea level pressure =====								
2015:1:2:0	1008,39	1008,37	1008,22	1008,14	1008,17	1008,13	1008,30	1008,35	1008,22

2015:1:3:0	1008,42	1008,39	1008,24	1008,26	1008,29	1008,23	1008,35	1008,41	1008,36
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	1007,35	1007,28	1007,11	1007,04	1007,01	1006,94	1007,07	1006,97	1006,94
2016:12:10:0	1008,27	1008,18	1007,81	1007,73	1007,70	1007,65	1007,96	1007,90	1007,78
	===== surface pressure =====								
2015:1:2:0	1008,44	1008,47	1008,22	1008,14	1007,42	1005,67	1002,30	1000,96	996,61
2015:1:3:0	1008,45	1008,46	1008,15	1008,17	1007,44	1005,62	1002,27	1000,94	996,54
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	1007,39	1007,36	1007,03	1007,02	1006,18	1004,33	1001,13	999,69	995,26
2016:12:10:0	1008,37	1008,41	1008,16	1008,11	1007,38	1005,66	1002,33	1001,02	996,68
	===== relative humidity 2m =====								
2015:1:2:0	77,76	77,38	77,07	76,39	70,32	70,08	68,62	69,77	71,50
2015:1:3:0	78,89	78,33	78,05	77,89	71,36	72,16	70,43	71,75	74,95
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	76,74	76,65	76,24	74,89	70,38	70,00	66,74	68,50	69,06

2016:12:10:0	77,28	76,56	76,32	75,73	68,26	67,96	65,78	66,93	69,03
	===== specific humidity 2m =====								
2015:1:2:0	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02
2015:1:3:0	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02
2016:12:10:0	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02
	===== temperature 2m =====								
2015:1:2:0	29,35	29,35	29,35	29,24	30,10	30,20	30,00	29,61	29,49
2015:1:3:0	29,39	29,41	29,42	29,28	30,21	30,09	29,81	29,87	29,15
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	30,50	30,34	30,27	30,45	32,03	32,88	32,38	31,79	32,09
2016:12:10:0	29,52	29,38	29,32	29,49	31,22	31,42	31,46	31,14	31,01
	===== u wind 10m =====								
2015:1:2:0	7,06	7,00	6,99	6,70	5,79	5,96	6,06	6,14	6,31
2015:1:3:0	7,50	7,42	7,65	7,33	6,32	6,34	6,75	6,70	6,56

:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	-0,31	-0,23	0,10	-0,66	-0,64	-0,19	-0,63	-0,99	-0,47
2016:12:10:0	6,19	6,11	6,12	5,86	5,06	5,18	5,14	5,24	5,29
	===== V wind 10m =====								
2015:1:2:0	-0,70	-0,57	-0,31	-0,44	-0,28	-0,06	-0,10	0,38	0,58
2015:1:3:0	0,85	0,95	1,16	0,87	0,90	1,07	0,61	0,99	1,19
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	-0,68	-0,30	0,46	-1,28	-0,72	0,31	-0,56	0,00	1,33
2016:12:10:0	-0,32	0,15	0,61	-0,48	0,14	0,52	-0,25	0,58	0,93
	===== surface vis =====								
2015:1:2:0	24160,83	24160,83	24160,83	24160,83	24160,83	24160,83	24160,83	24160,83	24160,83
2015:1:3:0	24135,36	24135,36	24135,36	24135,36	24043,70	23402,03	24135,36	23752,03	22727,03
:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:
2016:11:23:0	24107,61	24107,61	24107,61	24107,61	24107,61	24107,61	24107,61	24107,61	24107,61
2016:12:10:0	24207,94	24207,94	24207,94	24207,94	24207,94	24207,94	24207,94	24207,94	24207,94

## Lampiran 5. Output Reduksi PCA Stamet Juanda

### Principal Component Analysis: hujan\_total ; hujan\_total ; hujan\_total ; hujan\_t

#### Eigenanalysis of the Correlation Matrix

Eigenvalue	6,1505	1,5582	0,6622	0,2074	0,1885	0,1032	0,0823	0,0319
Proportion	0,683	0,173	0,074	0,023	0,021	0,011	0,009	0,004
Cumulative	0,683	0,857	0,930	0,953	0,974	0,986	0,995	0,998

Eigenvalue	0,0158
Proportion	0,002
Cumulative	1,000

Variable		PC1	PC2	PC3	PC4	PC5	PC6	PC7
hujan_total	_1_1	0,317	-0,353	0,391	-0,443	-0,253	0,571	-0,146
hujan_total	_0_1	0,313	-0,468	-0,132	0,146	-0,205	-0,295	-0,326
hujan_total	_-1_1	0,297	-0,417	-0,476	0,067	-0,239	-0,165	0,389
hujan_total	_1_0	0,326	-0,094	0,656	0,037	0,267	-0,424	0,444
hujan_total	_0_0	0,383	0,002	0,019	0,355	0,458	0,016	-0,590
hujan_total	_-1_0	0,361	0,056	-0,382	-0,208	0,578	0,345	0,315
hujan_total	_1_-1	0,341	0,344	0,097	0,516	-0,372	0,254	0,172
hujan_total	_0_-1	0,335	0,429	-0,063	0,037	-0,251	0,066	0,030
hujan_total	_-1_-1	0,320	0,410	-0,113	-0,583	-0,141	-0,438	-0,215

Variable		PC8	PC9
hujan_total	_1_1	0,121	-0,029
hujan_total	_0_1	-0,619	0,159
hujan_total	_-1_1	0,514	-0,086
hujan_total	_1_0	0,002	0,066
hujan_total	_0_0	0,405	-0,066
hujan_total	_-1_0	-0,352	-0,014
hujan_total	_1_-1	-0,195	-0,468
hujan_total	_0_-1	0,105	0,787
hujan_total	_-1_-1	0,019	-0,344

### Principal Component Analysis: dew2m\_1\_1; dew2m\_0\_1; dew2m\_-1\_1; dew2m\_1\_0;

#### Eigenanalysis of the Correlation Matrix

Eigenvalue	8,6105	0,2818	0,0688	0,0262	0,0071	0,0024	0,0015	0,0012
Proportion	0,957	0,031	0,008	0,003	0,001	0,000	0,000	0,000
Cumulative	0,957	0,988	0,996	0,999	0,999	1,000	1,000	1,000

Eigenvalue	0,0005
Proportion	0,000
Cumulative	1,000

Variable		PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
dew2m	_1_1	0,320	0,624	0,221	-0,337	0,408	0,057	-0,392	0,038
dew2m	_0_1	0,338	-0,016	0,426	0,257	0,305	-0,275	0,396	0,276
dew2m	_-1_1	0,336	-0,172	0,482	0,106	-0,072	0,328	0,220	-0,584
dew2m	_1_0	0,323	0,578	-0,247	0,028	-0,604	-0,004	0,335	-0,063
dew2m	_0_0	0,339	-0,113	0,019	0,344	-0,206	-0,538	-0,257	0,245
dew2m	_-1_0	0,336	-0,294	0,172	-0,076	-0,405	0,341	-0,527	0,225
dew2m	_1_-1	0,336	-0,034	-0,519	0,453	0,363	0,495	0,039	0,157
dew2m	_0_-1	0,337	-0,192	-0,385	-0,120	0,176	-0,404	-0,197	-0,593
dew2m	_-1_-1	0,333	-0,330	-0,170	-0,682	0,029	0,009	0,380	0,299

Variable		PC9
dew2m	_1_1	0,144
dew2m	_0_1	-0,486
dew2m	_-1_1	0,334
dew2m	_1_0	-0,136
dew2m	_0_0	0,544
dew2m	_-1_0	-0,395
dew2m	_1_-1	0,090
dew2m	_0_-1	-0,319



dew2m \_-1\_-1 0,226

### Principal Component Analysis: wind\_gust\_1; wind\_gust\_0; wind\_gust\_-; wind\_gu

#### Eigenanalysis of the Correlation Matrix

Eigenvalue	7,4931	0,8260	0,4607	0,1011	0,0443	0,0331	0,0271	0,0100
Proportion	0,833	0,092	0,051	0,011	0,005	0,004	0,003	0,001
Cumulative	0,833	0,924	0,976	0,987	0,992	0,995	0,998	0,999

Eigenvalue	0,0047
Proportion	0,001
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
wind_gust _1_1	0,343	-0,306	0,139	-0,432	0,063	0,440	0,426	-0,261
wind_gust _0_1	0,350	-0,093	0,363	-0,186	0,262	-0,462	-0,030	-0,407
wind_gust _-1_1	0,313	0,369	0,551	0,150	0,354	0,236	-0,337	0,295
wind_gust _1_0	0,329	-0,460	0,003	-0,207	-0,261	0,088	-0,171	0,647
wind_gust _0_0	0,360	-0,008	0,094	0,176	-0,409	-0,607	0,150	0,092
wind_gust _-1_0	0,318	0,517	0,016	0,061	-0,604	0,320	0,086	-0,211
wind_gust _1_-1	0,327	-0,398	-0,259	0,521	0,010	0,201	-0,450	-0,383
wind_gust _0_-1	0,347	0,113	-0,344	0,393	0,411	0,002	0,571	0,248
wind_gust _-1_-1	0,309	0,333	-0,593	-0,503	0,178	-0,136	-0,342	-0,001

Variable	PC9
wind_gust _1_1	0,368
wind_gust _0_1	-0,504
wind_gust _-1_1	0,240
wind_gust _1_0	-0,337
wind_gust _0_0	0,514
wind_gust _-1_0	-0,328
wind_gust _1_-1	0,079
wind_gust _0_-1	-0,194
wind_gust _-1_-1	0,151

### Principal Component Analysis: hi\_cloud\_1; hi\_cloud\_0; hi\_cloud\_-1; hi\_clou

#### Eigenanalysis of the Correlation Matrix

Eigenvalue	8,2272	0,4005	0,2129	0,0760	0,0403	0,0181	0,0124	0,0072
Proportion	0,914	0,045	0,024	0,008	0,004	0,002	0,001	0,001
Cumulative	0,914	0,959	0,982	0,991	0,995	0,997	0,999	0,999

Eigenvalue	0,0053
Proportion	0,001
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
hi_cloud _1_1	0,329	0,437	0,086	-0,596	0,090	-0,168	-0,369	-0,001
hi_cloud _0_1	0,336	0,365	-0,212	-0,137	-0,235	-0,265	0,257	-0,291
hi_cloud _-1_1	0,336	0,184	-0,425	0,212	-0,508	0,332	0,129	0,392
hi_cloud _1_0	0,334	0,295	0,402	0,131	0,442	0,537	0,088	0,238
hi_cloud _0_0	0,343	0,077	0,007	0,474	0,356	-0,468	0,365	-0,139
hi_cloud _-1_0	0,339	-0,189	-0,327	0,348	0,163	0,032	-0,726	-0,188
hi_cloud _1_-1	0,326	-0,225	0,649	0,128	-0,536	0,045	-0,079	-0,318
hi_cloud _0_-1	0,331	-0,468	0,098	-0,191	0,011	-0,388	0,010	0,653
hi_cloud _-1_-1	0,325	-0,495	-0,260	-0,406	0,205	0,361	0,326	-0,351

Variable	PC9
hi_cloud _1_1	-0,408
hi_cloud _0_1	0,643
hi_cloud _-1_1	-0,298
hi_cloud _1_0	0,274
hi_cloud _0_0	-0,391
hi_cloud _-1_0	0,175
hi_cloud _1_-1	-0,092
hi_cloud _0_-1	0,220

```
hi_cloud _-1_-1 -0,126
```

### Principal Component Analysis: low\_cloud\_1; low\_cloud\_0; low\_cloud\_-; low\_clo

Eigenanalysis of the Correlation Matrix

Eigenvalue	5,6407	1,7036	0,6375	0,4468	0,2428	0,1331	0,0845	0,0648
Proportion	0,627	0,189	0,071	0,050	0,027	0,015	0,009	0,007
Cumulative	0,627	0,816	0,887	0,937	0,963	0,978	0,988	0,995

Eigenvalue	0,0462
Proportion	0,005
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
low_cloud _1_1	0,288	0,451	-0,379	-0,260	0,196	0,595	-0,011	-0,330
low_cloud _0_1	0,340	0,386	0,131	-0,272	-0,198	-0,007	-0,155	0,761
low_cloud _-1_1	0,331	0,213	0,607	-0,203	-0,343	-0,176	0,190	-0,477
low_cloud _1_0	0,339	0,295	-0,420	0,170	0,185	-0,686	0,212	-0,094
low_cloud _0_0	0,386	-0,092	0,154	0,356	0,451	0,029	-0,134	0,103
low_cloud _-1_0	0,366	-0,230	0,365	0,085	0,402	0,113	-0,208	-0,053
low_cloud _1_-1	0,355	-0,158	-0,257	0,455	-0,623	0,118	-0,396	-0,116
low_cloud _0_-1	0,337	-0,410	-0,116	0,004	-0,129	0,245	0,766	0,201
low_cloud _-1_-1	0,231	-0,511	-0,244	-0,668	0,021	-0,238	-0,299	-0,091

Variable	PC9
low_cloud _1_1	0,011
low_cloud _0_1	-0,042
low_cloud _-1_1	0,155
low_cloud _1_0	-0,185
low_cloud _0_0	0,678
low_cloud _-1_0	-0,672
low_cloud _1_-1	-0,056
low_cloud _0_-1	-0,028
low_cloud _-1_-1	0,156

### Principal Component Analysis: mid\_cloud\_1; mid\_cloud\_0; mid\_cloud\_-; mid\_clo

Eigenanalysis of the Correlation Matrix

Eigenvalue	8,1726	0,4528	0,1619	0,1190	0,0414	0,0246	0,0128	0,0097
Proportion	0,908	0,050	0,018	0,013	0,005	0,003	0,001	0,001
Cumulative	0,908	0,958	0,976	0,990	0,994	0,997	0,998	0,999

Eigenvalue	0,0051
Proportion	0,001
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
mid_cloud _1_1	0,325	-0,486	0,057	-0,479	0,017	-0,174	-0,162	-0,415
mid_cloud _0_1	0,336	-0,384	0,154	-0,010	-0,249	-0,280	0,333	-0,005
mid_cloud _-1_1	0,338	-0,213	0,223	0,421	-0,530	0,263	-0,138	0,352
mid_cloud _1_0	0,338	-0,255	-0,352	-0,176	0,423	0,474	-0,251	0,377
mid_cloud _0_0	0,343	0,007	-0,265	0,337	0,399	-0,328	0,535	0,153
mid_cloud _-1_0	0,340	0,143	0,142	0,529	0,265	0,087	-0,341	-0,580
mid_cloud _1_-1	0,325	0,390	-0,582	-0,165	-0,468	0,212	0,174	-0,286
mid_cloud _0_-1	0,333	0,418	0,027	-0,171	-0,042	-0,573	-0,483	0,344
mid_cloud _-1_-1	0,322	0,398	0,607	-0,337	0,167	0,330	0,338	0,042

Variable	PC9
mid_cloud _1_1	-0,443
mid_cloud _0_1	0,682
mid_cloud _-1_1	-0,347
mid_cloud _1_0	0,241
mid_cloud _0_0	-0,348
mid_cloud _-1_0	0,184
mid_cloud _1_-1	0,018
mid_cloud _0_-1	0,051

mid\_cloud \_-1\_-1 -0,045

### Principal Component Analysis: tot\_cloud\_1; tot\_cloud\_0; tot\_cloud\_-; tot\_clo

#### Eigenanalysis of the Correlation Matrix

Eigenvalue	8,2474	0,4436	0,1322	0,0991	0,0366	0,0185	0,0103	0,0078
Proportion	0,916	0,049	0,015	0,011	0,004	0,002	0,001	0,001
Cumulative	0,916	0,966	0,980	0,991	0,995	0,997	0,999	0,999

Eigenvalue	0,0045
Proportion	0,001
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
tot_cloud _1_1	0,328	-0,434	0,059	-0,511	-0,000	-0,325	0,316	-0,241
tot_cloud _0_1	0,335	-0,388	-0,167	-0,067	0,207	-0,133	-0,297	-0,205
tot_cloud _-1_1	0,336	-0,247	-0,404	0,306	0,482	0,261	-0,055	0,364
tot_cloud _1_0	0,337	-0,259	0,418	-0,094	-0,423	0,383	0,084	0,524
tot_cloud _0_0	0,343	-0,007	0,065	0,388	-0,457	-0,208	-0,566	-0,229
tot_cloud _-1_0	0,340	0,157	-0,312	0,391	-0,289	0,051	0,652	-0,255
tot_cloud _1_-1	0,331	0,260	0,650	0,156	0,484	0,184	0,062	-0,322
tot_cloud _0_-1	0,330	0,459	0,015	-0,072	0,107	-0,633	0,033	0,502
tot_cloud _-1_-1	0,321	0,480	-0,326	-0,547	-0,090	0,425	-0,226	-0,139

Variable	PC9
tot_cloud _1_1	0,420
tot_cloud _0_1	-0,717
tot_cloud _-1_1	0,365
tot_cloud _1_0	-0,169
tot_cloud _0_0	0,320
tot_cloud _-1_0	-0,181
tot_cloud _1_-1	0,020
tot_cloud _0_-1	-0,100
tot_cloud _-1_-1	0,045

### Principal Component Analysis: mslp\_1\_1; mslp\_0\_1; mslp\_-1\_1; mslp\_1\_0; mslp

#### Eigenanalysis of the Correlation Matrix

Eigenvalue	8,9487	0,0281	0,0131	0,0051	0,0021	0,0015	0,0010	0,0003
Proportion	0,994	0,003	0,001	0,001	0,000	0,000	0,000	0,000
Cumulative	0,994	0,997	0,999	0,999	1,000	1,000	1,000	1,000

Eigenvalue	0,0001
Proportion	0,000
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
mslp _1_1	0,333	0,302	-0,458	-0,134	0,599	-0,401	0,018	0,066
mslp _0_1	0,334	0,263	-0,266	0,170	-0,300	-0,114	-0,416	-0,082
mslp _-1_1	0,334	0,357	-0,009	0,319	-0,488	0,017	0,046	-0,210
mslp _1_0	0,333	-0,304	-0,379	-0,433	-0,397	0,147	0,220	0,486
mslp _0_0	0,334	-0,070	-0,010	-0,177	0,297	0,737	-0,379	-0,255
mslp _-1_0	0,334	0,291	0,240	0,200	0,179	0,290	0,684	0,097
mslp _1_-1	0,332	-0,621	-0,062	0,096	0,014	-0,259	0,250	-0,596
mslp _0_-1	0,333	-0,356	0,286	0,519	0,172	-0,106	-0,288	0,527
mslp _-1_-1	0,333	0,135	0,658	-0,561	-0,074	-0,314	-0,135	-0,035

Variable	PC9
mslp _1_1	-0,215
mslp _0_1	0,661
mslp _-1_1	-0,612
mslp _1_0	-0,045
mslp _0_0	-0,108
mslp _-1_0	0,335
mslp _1_-1	0,075
mslp _0_-1	-0,101

```
mslp _-1_-1 0,010
```

### Principal Component Analysis: sur\_press\_1; sur\_press\_0; sur\_press\_-; sur\_pre

Eigenanalysis of the Correlation Matrix

Eigenvalue	8,9181	0,0545	0,0208	0,0042	0,0014	0,0006	0,0003	0,0001
Proportion	0,991	0,006	0,002	0,000	0,000	0,000	0,000	0,000
Cumulative	0,991	0,997	0,999	1,000	1,000	1,000	1,000	1,000

Eigenvalue	0,0000
Proportion	0,000
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
sur_press_1_1	0,333	-0,403	-0,145	0,515	-0,122	-0,137	-0,557	-0,089
sur_press_0_1	0,334	-0,356	0,087	0,125	-0,290	0,031	0,203	0,497
sur_press_-1_1	0,334	-0,279	0,366	-0,267	-0,392	0,304	0,306	-0,403
sur_press_1_0	0,334	-0,214	-0,371	0,096	0,552	-0,058	0,386	-0,437
sur_press_0_0	0,335	-0,065	-0,089	-0,343	0,214	-0,383	0,189	0,521
sur_press_-1_0	0,334	0,026	0,361	-0,476	0,323	-0,020	-0,567	-0,104
sur_press_1_-1	0,333	0,316	-0,550	-0,145	-0,105	0,636	-0,154	0,157
sur_press_0_-1	0,332	0,511	-0,148	-0,023	-0,456	-0,549	0,039	-0,266
sur_press_-1_-1	0,332	0,470	0,488	0,523	0,272	0,177	0,153	0,123

Variable	PC9
sur_press_1_1	0,298
sur_press_0_1	-0,605
sur_press_-1_1	0,321
sur_press_1_0	-0,219
sur_press_0_0	0,509
sur_press_-1_0	-0,308
sur_press_1_-1	0,048
sur_press_0_-1	-0,159
sur_press_-1_-1	0,115

### Principal Component Analysis: rhum2m\_1\_1; rhum2m\_0\_1; rhum2m\_-1\_1; rhum2m\_1

Eigenanalysis of the Correlation Matrix

Eigenvalue	7,9244	0,7707	0,2090	0,0457	0,0237	0,0137	0,0060	0,0046
Proportion	0,880	0,086	0,023	0,005	0,003	0,002	0,001	0,001
Cumulative	0,880	0,966	0,989	0,994	0,997	0,999	0,999	1,000

Eigenvalue	0,0021
Proportion	0,000
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
rhum2m_1_1	0,284	0,650	-0,362	-0,134	-0,472	0,039	-0,075	0,327
rhum2m_0_1	0,348	-0,110	-0,340	0,319	-0,197	-0,308	0,171	-0,346
rhum2m_-1_1	0,344	-0,185	-0,398	0,192	0,076	0,320	-0,149	-0,437
rhum2m_1_0	0,289	0,634	0,327	0,015	0,521	-0,002	0,039	-0,358
rhum2m_0_0	0,351	-0,114	0,024	0,303	0,280	-0,585	0,122	0,466
rhum2m_-1_0	0,348	-0,182	-0,177	-0,062	0,412	0,491	-0,185	0,455
rhum2m_1_-1	0,340	-0,092	0,557	0,308	-0,394	0,396	0,385	0,089
rhum2m_0_-1	0,345	-0,167	0,382	-0,183	-0,242	-0,218	-0,745	-0,098
rhum2m_-1_-1	0,343	-0,215	-0,014	-0,786	0,006	-0,112	0,434	-0,109

Variable	PC9
rhum2m_1_1	0,103
rhum2m_0_1	-0,605
rhum2m_-1_1	0,575
rhum2m_1_0	-0,077
rhum2m_0_0	0,344
rhum2m_-1_0	-0,397
rhum2m_1_-1	0,047
rhum2m_0_-1	-0,048

rhum2m \_-1\_-1 0,070

### Principal Component Analysis: shum2m \_1\_1; shum2m \_0\_1; shum2m \_-1\_1; shum2m \_1

#### Eigenanalysis of the Correlation Matrix

Eigenvalue	8,6105	0,2820	0,0688	0,0252	0,0074	0,0025	0,0018	0,0013
Proportion	0,957	0,031	0,008	0,003	0,001	0,000	0,000	0,000
Cumulative	0,957	0,988	0,996	0,998	0,999	1,000	1,000	1,000

Eigenvalue	0,0006
Proportion	0,000
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
shum2m _1_1	0,320	0,633	0,212	-0,309	-0,410	0,093	-0,397	0,032
shum2m _0_1	0,338	-0,011	0,430	0,268	-0,304	-0,279	0,369	0,285
shum2m _-1_1	0,336	-0,168	0,482	0,098	0,078	0,307	0,254	-0,577
shum2m _1_0	0,323	0,573	-0,254	0,018	0,603	-0,036	0,341	-0,056
shum2m _0_0	0,339	-0,120	0,020	0,337	0,204	-0,538	-0,298	0,235
shum2m _-1_0	0,336	-0,297	0,174	-0,100	0,419	0,358	-0,509	0,197
shum2m _1_-1	0,336	-0,051	-0,514	0,463	-0,329	0,506	0,061	0,170
shum2m _0_-1	0,337	-0,189	-0,390	-0,114	-0,202	-0,379	-0,194	-0,600
shum2m _-1_-1	0,333	-0,318	-0,163	-0,686	-0,056	-0,021	0,370	0,314

Variable	PC9
shum2m _1_1	-0,143
shum2m _0_1	0,491
shum2m _-1_1	-0,345
shum2m _1_0	0,132
shum2m _0_0	-0,530
shum2m _-1_0	0,396
shum2m _1_-1	-0,097
shum2m _0_-1	0,322
shum2m _-1_-1	-0,226

### Principal Component Analysis: temp2m \_1\_1; temp2m \_0\_1; temp2m \_-1\_1; temp2m \_1

#### Eigenanalysis of the Correlation Matrix

Eigenvalue	6,5622	1,9501	0,2609	0,1093	0,0435	0,0374	0,0306	0,0060
Proportion	0,729	0,217	0,029	0,012	0,005	0,004	0,003	0,001
Cumulative	0,729	0,946	0,975	0,987	0,992	0,996	0,999	1,000

Eigenvalue	0,0000
Proportion	0,000
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
temp2m _1_1	-0,088	0,697	0,008	0,070	0,002	0,013	0,023	-0,018
temp2m _0_1	0,380	0,072	0,223	-0,120	0,739	-0,004	-0,123	0,474
temp2m _-1_1	0,381	0,032	0,371	-0,047	0,181	-0,292	0,041	-0,771
temp2m _1_0	-0,088	0,697	0,008	0,070	0,002	0,013	0,023	-0,018
temp2m _0_0	0,377	0,058	0,323	-0,204	-0,353	0,732	-0,218	-0,015
temp2m _-1_0	0,380	0,066	0,281	-0,051	-0,529	-0,525	0,206	0,415
temp2m _1_-1	0,368	0,063	-0,538	-0,357	0,058	0,161	0,640	-0,066
temp2m _0_-1	0,367	0,086	-0,569	-0,075	-0,112	-0,224	-0,680	-0,063
temp2m _-1_-1	0,371	-0,049	-0,125	0,892	0,011	0,165	0,143	0,015

Variable	PC9
temp2m _1_1	-0,707
temp2m _0_1	-0,000
temp2m _-1_1	0,000
temp2m _1_0	0,707
temp2m _0_0	0,000
temp2m _-1_0	-0,000
temp2m _1_-1	0,000
temp2m _0_-1	-0,000

temp2m \_-1\_-1 0,000

### Principal Component Analysis: uwind10m \_1\_; uwind10m \_0\_; uwind10m \_-1\_; uwind10

Eigenanalysis of the Correlation Matrix

Eigenvalue	8,8502	0,0784	0,0458	0,0130	0,0052	0,0040	0,0022	0,0009
Proportion	0,983	0,009	0,005	0,001	0,001	0,000	0,000	0,000
Cumulative	0,983	0,992	0,997	0,999	0,999	1,000	1,000	1,000

Eigenvalue	0,0003
Proportion	0,000
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
uwind10m _1_1	0,333	-0,400	0,179	-0,443	-0,111	-0,072	-0,453	-0,339
uwind10m _0_1	0,334	-0,287	0,257	0,167	-0,537	-0,112	0,159	-0,147
uwind10m _-1_1	0,334	-0,006	0,409	0,481	-0,075	0,337	-0,147	0,481
uwind10m _1_0	0,334	-0,355	-0,133	-0,407	0,436	0,218	0,122	0,469
uwind10m _0_0	0,336	-0,104	0,014	0,110	0,182	-0,483	0,665	-0,064
uwind10m _-1_0	0,333	0,360	0,312	0,161	0,585	0,021	-0,172	-0,435
uwind10m _1_-1	0,332	-0,114	-0,697	0,331	-0,052	0,411	0,009	-0,322
uwind10m _0_-1	0,334	0,282	-0,366	0,072	-0,090	-0,586	-0,442	0,339
uwind10m _-1_-1	0,330	0,631	0,021	-0,476	-0,342	0,274	0,258	0,015

Variable	PC9
uwind10m _1_1	0,404
uwind10m _0_1	-0,604
uwind10m _-1_1	0,343
uwind10m _1_0	-0,326
uwind10m _0_0	0,390
uwind10m _-1_0	-0,273
uwind10m _1_-1	0,082
uwind10m _0_-1	-0,095
uwind10m _-1_-1	0,079

### Principal Component Analysis: vwind10m \_1\_; vwind10m \_0\_; vwind10m \_-1\_; vwind10

Eigenanalysis of the Correlation Matrix

Eigenvalue	6,4967	1,9287	0,3957	0,0880	0,0545	0,0228	0,0084	0,0038
Proportion	0,722	0,214	0,044	0,010	0,006	0,003	0,001	0,000
Cumulative	0,722	0,936	0,980	0,990	0,996	0,998	0,999	1,000

Eigenvalue	0,0014
Proportion	0,000
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
vwind10m _1_1	0,340	0,328	-0,213	-0,436	0,166	0,189	-0,537	0,005
vwind10m _0_1	0,355	0,293	0,126	-0,025	-0,240	0,379	-0,062	-0,443
vwind10m _-1_1	0,359	0,078	0,570	-0,120	-0,571	-0,171	0,128	0,289
vwind10m _1_0	0,353	0,286	-0,234	-0,153	0,305	-0,163	0,354	0,574
vwind10m _0_0	0,381	0,037	0,203	0,516	0,418	0,356	0,315	-0,109
vwind10m _-1_0	0,272	-0,466	0,450	-0,226	0,471	-0,319	-0,253	-0,141
vwind10m _1_-1	0,375	0,073	-0,388	0,258	-0,149	-0,657	0,035	-0,413
vwind10m _0_-1	0,299	-0,439	-0,247	0,429	-0,229	0,186	-0,449	0,403
vwind10m _-1_-1	0,233	-0,550	-0,323	-0,452	-0,159	0,265	0,447	-0,159

Variable	PC9
vwind10m _1_1	0,435
vwind10m _0_1	-0,608
vwind10m _-1_1	0,265
vwind10m _1_0	-0,375
vwind10m _0_0	0,366
vwind10m _-1_0	-0,219
vwind10m _1_-1	0,110
vwind10m _0_-1	-0,147

```
vwindl0m _-1_-1 0,121
```

# **Principal Component Analysis: sur\_vis\_1\_1; sur\_vis\_0\_1; sur\_vis\_-1\_; sur\_vis**

## Eigenanalysis of the Correlation Matrix

Eigenvalue	6,6217	1,4003	0,4501	0,2576	0,1186	0,0831	0,0340	0,0197
Proportion	0,736	0,156	0,050	0,029	0,013	0,009	0,004	0,002
Cumulative	0,736	0,891	0,941	0,970	0,983	0,992	0,996	0,998

Eigenvalue	0,0148
Proportion	0,002
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
sur_vis_1_1	0,294	-0,468	0,336	-0,477	-0,139	-0,133	0,436	0,021
sur_vis_0_1	0,310	-0,490	-0,025	-0,187	0,001	-0,194	-0,515	0,038
sur_vis_-1_1	0,323	-0,321	-0,500	0,193	0,582	0,181	-0,015	-0,079
sur_vis_1_0	0,336	-0,094	0,587	0,508	0,037	0,475	0,073	-0,128
sur_vis_0_0	0,371	0,068	-0,070	0,344	-0,551	-0,239	-0,381	0,101
sur_vis_-1_0	0,363	0,081	-0,426	0,206	-0,270	-0,150	0,606	0,007
sur_vis_1_-1	0,324	0,395	0,284	0,027	0,491	-0,459	0,028	0,453
sur_vis_0_-1	0,334	0,401	0,040	-0,278	0,081	-0,108	-0,107	-0,787
sur_vis_-1_-1	0,338	0,320	-0,153	-0,452	-0,128	0,620	-0,118	0,376

Variable	PC9
sur_vis_1_1	-0,356
sur_vis_0_1	0,569
sur_vis_-1_1	-0,358
sur_vis_1_0	0,165
sur_vis_0_0	-0,467
sur_vis_-1_0	0,419
sur_vis_1_-1	0,022
sur_vis_0_-1	-0,009
sur_vis_-1_-1	0,016

## Lampiran 6. Output Reduksi PCA Stamet Ngurai Rai

### Principal Component Analysis: hujan\_total ; hujan\_total ; hujan\_total ; hujan\_t

#### Eigenanalysis of the Correlation Matrix

Eigenvalue	7,5120	0,8136	0,3482	0,1525	0,0726	0,0491	0,0247	0,0163
Proportion	0,835	0,090	0,039	0,017	0,008	0,005	0,003	0,002
Cumulative	0,835	0,925	0,964	0,981	0,989	0,994	0,997	0,999

Eigenvalue	0,0110
Proportion	0,001
Cumulative	1,000

Variable		PC1	PC2	PC3	PC4	PC5	PC6	PC7
hujan_total	_1_1	0,317	-0,492	0,079	-0,304	-0,569	-0,434	0,186
hujan_total	_0_1	0,296	-0,615	0,002	-0,243	0,313	0,564	-0,137
hujan_total	_-1_1	0,337	-0,282	0,018	0,626	0,459	-0,344	0,156
hujan_total	_1_0	0,320	0,231	-0,691	-0,257	0,195	-0,145	0,386
hujan_total	_0_0	0,345	0,132	-0,451	0,143	-0,296	0,302	-0,365
hujan_total	_-1_0	0,353	0,102	0,066	0,468	-0,367	0,040	-0,225
hujan_total	_1_-1	0,344	0,258	0,240	-0,341	0,312	-0,309	-0,400
hujan_total	_0_-1	0,346	0,260	0,302	-0,184	0,066	-0,046	-0,184
hujan_total	_-1_-1	0,338	0,292	0,399	0,019	-0,083	0,405	0,629

Variable		PC8	PC9
hujan_total	_1_1	-0,105	-0,044
hujan_total	_0_1	0,193	0,048
hujan_total	_-1_1	-0,249	-0,010
hujan_total	_1_0	0,274	0,135
hujan_total	_0_0	-0,539	-0,192
hujan_total	_-1_0	0,670	0,073
hujan_total	_1_-1	0,105	-0,526
hujan_total	_0_-1	-0,229	0,771
hujan_total	_-1_-1	-0,106	-0,252

### Principal Component Analysis: dew2m\_1\_1; dew2m\_0\_1; dew2m\_-1\_1; dew2m\_1\_0;

#### Eigenanalysis of the Correlation Matrix

Eigenvalue	8,4738	0,4583	0,0323	0,0196	0,0102	0,0044	0,0010	0,0004
Proportion	0,942	0,051	0,004	0,002	0,001	0,000	0,000	0,000
Cumulative	0,942	0,992	0,996	0,998	0,999	1,000	1,000	1,000

Eigenvalue	0,0001
Proportion	0,000
Cumulative	1,000

Variable		PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
dew2m	_1_1	0,316	-0,568	-0,297	0,345	0,527	-0,270	0,107	-0,079
dew2m	_0_1	0,317	-0,555	-0,238	-0,367	-0,423	0,454	-0,027	0,105
dew2m	_-1_1	0,338	-0,203	0,599	-0,356	-0,137	-0,574	-0,089	-0,064
dew2m	_1_0	0,342	0,058	0,203	0,586	-0,177	0,108	-0,515	0,332
dew2m	_0_0	0,342	0,106	0,275	0,208	-0,004	0,399	0,025	-0,621
dew2m	_-1_0	0,340	0,179	0,320	-0,117	0,376	0,321	0,541	0,398
dew2m	_1_-1	0,336	0,271	-0,317	0,204	-0,468	-0,326	0,413	0,228
dew2m	_0_-1	0,335	0,309	-0,297	-0,112	-0,045	-0,109	0,045	-0,486
dew2m	_-1_-1	0,333	0,336	-0,297	-0,406	0,361	-0,003	-0,498	0,190

Variable		PC9
dew2m	_1_1	0,044
dew2m	_0_1	-0,055
dew2m	_-1_1	0,013
dew2m	_1_0	-0,275
dew2m	_0_0	0,456
dew2m	_-1_0	-0,202
dew2m	_1_-1	0,352
dew2m	_0_-1	-0,663



dew2m \_-1\_-1 0,329

### Principal Component Analysis: wind\_gust\_1; wind\_gust\_0; wind\_gust\_-; wind\_gu

#### Eigenanalysis of the Correlation Matrix

Eigenvalue	8,6571	0,1738	0,1128	0,0272	0,0163	0,0087	0,0025	0,0012
Proportion	0,962	0,019	0,013	0,003	0,002	0,001	0,000	0,000
Cumulative	0,962	0,981	0,994	0,997	0,999	1,000	1,000	1,000

Eigenvalue	0,0004
Proportion	0,000
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
wind_gust _1 1	0,323	-0,686	0,338	0,106	0,164	-0,453	0,119	-0,231
wind_gust _0 1	0,331	-0,457	-0,247	0,353	-0,307	0,572	-0,023	0,268
wind_gust _-1 1	0,328	-0,088	-0,725	-0,450	-0,116	-0,301	-0,185	-0,132
wind_gust _1 0	0,336	-0,017	0,303	-0,492	0,390	0,175	-0,302	0,517
wind_gust _0 0	0,338	0,186	-0,036	0,136	0,472	0,388	-0,148	-0,569
wind_gust _-1 0	0,336	0,274	-0,246	0,229	0,364	-0,131	0,645	0,244
wind_gust _1 -1	0,336	0,162	0,301	-0,390	-0,478	0,129	0,489	-0,086
wind_gust _0 -1	0,337	0,265	0,217	0,069	-0,308	0,003	-0,236	-0,323
wind_gust _-1 -1	0,335	0,326	0,087	0,439	-0,185	-0,402	-0,358	0,308

Variable	PC9
wind_gust _1 1	-0,003
wind_gust _0 1	0,033
wind_gust _-1 1	-0,028
wind_gust _1 0	0,105
wind_gust _0 0	-0,335
wind_gust _-1 0	0,273
wind_gust _1 -1	-0,356
wind_gust _0 -1	0,715
wind_gust _-1 -1	-0,404

### Principal Component Analysis: hi\_cloud\_1; hi\_cloud\_0; hi\_cloud\_-1; hi\_clou

#### Eigenanalysis of the Correlation Matrix

Eigenvalue	8,2820	0,3365	0,2123	0,0678	0,0600	0,0182	0,0122	0,0064
Proportion	0,920	0,037	0,024	0,008	0,007	0,002	0,001	0,001
Cumulative	0,920	0,958	0,981	0,989	0,995	0,997	0,999	0,999

Eigenvalue	0,0047
Proportion	0,001
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
hi_cloud _1 1	0,329	-0,258	-0,526	-0,460	0,292	0,063	-0,205	0,382
hi_cloud _0 1	0,330	-0,506	-0,123	0,022	0,196	-0,357	0,159	-0,521
hi_cloud _-1 1	0,324	-0,517	0,302	0,546	-0,011	0,293	0,079	0,283
hi_cloud _1 0	0,335	0,181	-0,423	0,064	-0,436	0,580	0,209	-0,139
hi_cloud _0 0	0,343	0,076	0,062	-0,193	-0,511	-0,392	0,272	-0,153
hi_cloud _-1 0	0,337	-0,040	0,426	-0,274	-0,362	-0,069	-0,546	0,188
hi_cloud _1 -1	0,331	0,404	-0,246	0,534	0,152	-0,223	-0,515	-0,156
hi_cloud _0 -1	0,337	0,374	0,096	0,072	0,245	-0,279	0,494	0,503
hi_cloud _-1 -1	0,333	0,259	0,427	-0,289	0,460	0,401	0,042	-0,384

Variable	PC9
hi_cloud _1 1	-0,244
hi_cloud _0 1	0,396
hi_cloud _-1 1	-0,258
hi_cloud _1 0	0,288
hi_cloud _0 0	-0,569
hi_cloud _-1 0	0,399
hi_cloud _1 -1	-0,142
hi_cloud _0 -1	0,312

```
hi_cloud _-1_-1 -0,185
```

### Principal Component Analysis: low\_cloud\_1; low\_cloud\_0; low\_cloud\_-; low\_clo

#### Eigenanalysis of the Correlation Matrix

Eigenvalue	7,3266	0,8689	0,3737	0,1649	0,1377	0,0606	0,0349	0,0199
Proportion	0,814	0,097	0,042	0,018	0,015	0,007	0,004	0,002
Cumulative	0,814	0,911	0,952	0,970	0,986	0,992	0,996	0,999

Eigenvalue	0,0128
Proportion	0,001
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
low_cloud_1_1	0,320	-0,250	-0,680	-0,165	-0,058	-0,508	-0,263	-0,039
low_cloud_0_1	0,303	-0,561	-0,196	-0,005	-0,308	0,625	0,191	0,039
low_cloud_-1_1	0,276	-0,573	0,622	0,058	0,051	-0,406	0,090	-0,084
low_cloud_1_0	0,352	0,153	-0,151	0,469	0,320	-0,204	0,478	0,167
low_cloud_0_0	0,357	0,130	-0,091	0,211	0,433	0,319	0,025	-0,351
low_cloud_-1_0	0,358	0,029	0,192	-0,142	0,392	0,199	-0,650	0,332
low_cloud_1_-1	0,334	0,305	0,129	0,472	-0,636	-0,040	-0,240	0,226
low_cloud_0_-1	0,348	0,294	0,144	-0,274	-0,216	-0,026	-0,019	-0,709
low_cloud_-1_-1	0,342	0,273	0,086	-0,622	-0,065	-0,019	0,420	0,417

Variable	PC9
low_cloud_1_1	-0,114
low_cloud_0_1	0,175
low_cloud_-1_1	-0,154
low_cloud_1_0	0,457
low_cloud_0_0	-0,624
low_cloud_-1_0	0,296
low_cloud_1_-1	-0,204
low_cloud_0_-1	0,382
low_cloud_-1_-1	-0,242

### Principal Component Analysis: mid\_cloud\_1; mid\_cloud\_0; mid\_cloud\_-; mid\_clo

#### Eigenanalysis of the Correlation Matrix

Eigenvalue	8,4299	0,2934	0,1511	0,0595	0,0378	0,0118	0,0089	0,0051
Proportion	0,937	0,033	0,017	0,007	0,004	0,001	0,001	0,001
Cumulative	0,937	0,969	0,986	0,993	0,997	0,998	0,999	1,000

Eigenvalue	0,0024
Proportion	0,000
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
mid_cloud_1_1	0,328	-0,427	-0,389	-0,177	0,629	0,188	0,251	-0,154
mid_cloud_0_1	0,332	-0,456	0,057	-0,241	-0,079	-0,228	-0,594	0,397
mid_cloud_-1_1	0,329	-0,409	0,448	-0,132	-0,462	0,009	0,307	-0,398
mid_cloud_1_0	0,336	0,028	-0,452	0,399	-0,242	-0,539	0,268	-0,076
mid_cloud_0_0	0,340	0,061	-0,083	0,564	-0,012	0,175	-0,437	-0,087
mid_cloud_-1_0	0,336	0,057	0,449	0,406	0,196	0,349	0,260	0,368
mid_cloud_1_-1	0,334	0,303	-0,335	-0,374	-0,386	0,323	0,216	0,445
mid_cloud_0_-1	0,335	0,395	-0,032	-0,226	-0,007	0,260	-0,329	-0,557
mid_cloud_-1_-1	0,329	0,433	0,344	-0,245	0,371	-0,545	0,071	0,056

Variable	PC9
mid_cloud_1_1	0,093
mid_cloud_0_1	-0,229
mid_cloud_-1_1	0,205
mid_cloud_1_0	-0,309
mid_cloud_0_0	0,571
mid_cloud_-1_0	-0,391
mid_cloud_1_-1	0,215
mid_cloud_0_-1	-0,440

mid\_cloud \_-1\_-1 0,287

### Principal Component Analysis: tot\_cloud\_1; tot\_cloud\_0; tot\_cloud\_-; tot\_clo

#### Eigenanalysis of the Correlation Matrix

Eigenvalue	8,3938	0,3020	0,1696	0,0541	0,0500	0,0130	0,0089	0,0051
Proportion	0,933	0,034	0,019	0,006	0,006	0,001	0,001	0,001
Cumulative	0,933	0,966	0,985	0,991	0,997	0,998	0,999	1,000

Eigenvalue	0,0035
Proportion	0,000
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
tot_cloud _1_1	0,329	-0,349	-0,442	0,063	-0,595	0,105	0,405	0,206
tot_cloud _0_1	0,330	-0,504	-0,035	-0,173	-0,090	-0,368	-0,545	-0,399
tot_cloud _-1_1	0,327	-0,463	0,355	-0,334	0,445	0,297	0,175	0,337
tot_cloud _1_0	0,336	0,118	-0,441	0,297	0,332	0,561	-0,290	-0,106
tot_cloud _0_0	0,341	0,078	0,026	0,558	0,150	-0,340	-0,148	0,314
tot_cloud _-1_0	0,337	0,020	0,460	0,424	0,006	-0,111	0,426	-0,374
tot_cloud _1_-1	0,332	0,344	-0,347	-0,421	0,297	-0,251	0,377	-0,344
tot_cloud _0_-1	0,336	0,378	0,040	-0,248	-0,094	-0,295	-0,174	0,536
tot_cloud _-1_-1	0,331	0,358	0,385	-0,192	-0,458	0,416	-0,221	-0,174

Variable	PC9
tot_cloud _1_1	-0,020
tot_cloud _0_1	0,077
tot_cloud _-1_1	-0,102
tot_cloud _1_0	0,264
tot_cloud _0_0	-0,555
tot_cloud _-1_0	0,402
tot_cloud _1_-1	-0,248
tot_cloud _0_-1	0,518
tot_cloud _-1_-1	-0,339

### Principal Component Analysis: mslp\_1\_1; mslp\_0\_1; mslp\_-1\_1; mslp\_1\_0; mslp

#### Eigenanalysis of the Correlation Matrix

Eigenvalue	8,9588	0,0223	0,0097	0,0065	0,0014	0,0009	0,0003	0,0001
Proportion	0,995	0,002	0,001	0,001	0,000	0,000	0,000	0,000
Cumulative	0,995	0,998	0,999	1,000	1,000	1,000	1,000	1,000

Eigenvalue	0,0000
Proportion	0,000
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
mslp_1_1	0,333	-0,465	-0,218	0,112	-0,489	-0,395	-0,441	0,078
mslp_0_1	0,333	-0,504	0,088	0,053	-0,183	0,701	0,278	-0,054
mslp_-1_1	0,333	-0,240	0,746	-0,020	0,354	-0,337	0,091	0,139
mslp_1_0	0,334	-0,150	-0,484	-0,223	0,295	-0,362	0,480	-0,317
mslp_0_0	0,334	0,088	-0,282	-0,435	0,342	0,264	-0,241	0,507
mslp_-1_0	0,334	0,276	0,200	-0,377	-0,042	0,120	-0,477	-0,546
mslp_1_-1	0,333	0,179	-0,153	0,756	0,380	0,124	-0,194	-0,201
mslp_0_-1	0,334	0,371	-0,002	0,174	-0,198	-0,089	0,113	0,511
mslp_-1_-1	0,333	0,443	0,106	-0,038	-0,460	-0,025	0,391	-0,118

Variable	PC9
mslp_1_1	-0,127
mslp_0_1	0,141
mslp_-1_1	-0,092
mslp_1_0	0,183
mslp_0_0	-0,331
mslp_-1_0	0,298
mslp_1_-1	-0,155
mslp_0_-1	0,632

```
mslp _-1_-1 -0,548
```

### Principal Component Analysis: sur\_press\_1; sur\_press\_0; sur\_press\_-; sur\_pre

#### Eigenanalysis of the Correlation Matrix

Eigenvalue	8,9638	0,0256	0,0074	0,0017	0,0009	0,0003	0,0001	0,0000
Proportion	0,996	0,003	0,001	0,000	0,000	0,000	0,000	0,000
Cumulative	0,996	0,999	1,000	1,000	1,000	1,000	1,000	1,000

Eigenvalue	0,0000
Proportion	0,000
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
sur_press_1_1	0,333	-0,383	-0,635	-0,124	-0,417	-0,371	-0,034	-0,092
sur_press_0_1	0,333	-0,521	-0,051	-0,259	0,273	0,654	-0,078	0,173
sur_press_-1_1	0,333	-0,431	0,516	0,076	0,335	-0,505	0,167	-0,162
sur_press_1_0	0,334	0,083	-0,127	0,656	-0,042	0,096	0,494	0,324
sur_press_0_0	0,334	0,100	0,154	0,315	-0,300	0,368	-0,104	-0,498
sur_press_-1_0	0,334	0,096	0,413	-0,016	-0,442	-0,081	-0,535	0,313
sur_press_1_-1	0,333	0,342	-0,304	0,134	0,544	-0,167	-0,449	0,228
sur_press_0_-1	0,333	0,353	-0,093	-0,224	0,190	0,017	0,074	-0,586
sur_press_-1_-1	0,333	0,358	0,126	-0,560	-0,140	-0,013	0,465	0,298

Variable	PC9
sur_press_1_1	0,051
sur_press_0_1	-0,096
sur_press_-1_1	0,098
sur_press_1_0	-0,273
sur_press_0_0	0,521
sur_press_-1_0	-0,350
sur_press_1_-1	0,289
sur_press_0_-1	-0,566
sur_press_-1_-1	0,326

### Principal Component Analysis: rhum2m\_1\_1; rhum2m\_0\_1; rhum2m\_-1\_1; rhum2m\_1

#### Eigenanalysis of the Correlation Matrix

Eigenvalue	6,9746	1,7712	0,1554	0,0626	0,0245	0,0086	0,0021	0,0008
Proportion	0,775	0,197	0,017	0,007	0,003	0,001	0,000	0,000
Cumulative	0,775	0,972	0,989	0,996	0,999	1,000	1,000	1,000

Eigenvalue	0,0002
Proportion	0,000
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
rhum2m_1_1	0,191	-0,631	-0,308	0,593	-0,217	-0,260	-0,042	0,029
rhum2m_0_1	0,127	-0,698	-0,042	-0,596	0,165	0,330	0,012	-0,049
rhum2m_-1_1	0,352	-0,143	0,794	-0,101	-0,149	-0,435	0,016	0,048
rhum2m_1_0	0,375	0,046	0,090	0,311	0,553	0,172	0,557	-0,252
rhum2m_0_0	0,375	0,076	0,102	0,193	0,081	0,474	-0,182	0,623
rhum2m_-1_0	0,372	0,126	0,077	0,102	-0,399	0,429	-0,393	-0,523
rhum2m_1_-1	0,369	0,134	-0,270	-0,152	0,457	-0,388	-0,453	-0,260
rhum2m_0_-1	0,368	0,153	-0,283	-0,204	-0,035	-0,200	-0,062	0,442
rhum2m_-1_-1	0,364	0,169	-0,308	-0,271	-0,471	-0,068	0,539	-0,076

Variable	PC9
rhum2m_1_1	0,022
rhum2m_0_1	-0,031
rhum2m_-1_1	0,023
rhum2m_1_0	-0,208
rhum2m_0_0	0,391
rhum2m_-1_0	-0,240
rhum2m_1_-1	0,345
rhum2m_0_-1	-0,692

rhsm2m \_-1\_-1 0,383

### Principal Component Analysis: shum2m \_1\_1; shum2m \_0\_1; shum2m \_-1\_1; shum2m \_1

#### Eigenanalysis of the Correlation Matrix

Eigenvalue	8,4812	0,4502	0,0327	0,0194	0,0103	0,0045	0,0011	0,0005
Proportion	0,942	0,050	0,004	0,002	0,001	0,001	0,000	0,000
Cumulative	0,942	0,992	0,996	0,998	0,999	1,000	1,000	1,000

Eigenvalue	0,0001
Proportion	0,000
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
shum2m _1_1	0,316	-0,569	-0,298	0,339	0,528	-0,273	0,108	-0,079
shum2m _0_1	0,318	-0,549	-0,245	-0,355	-0,430	0,460	-0,025	0,103
shum2m _-1_1	0,337	-0,207	0,591	-0,371	-0,137	-0,571	-0,092	-0,056
shum2m _1_0	0,342	0,056	0,212	0,587	-0,169	0,108	-0,511	0,348
shum2m _0_0	0,342	0,103	0,278	0,209	-0,005	0,382	0,017	-0,641
shum2m _-1_0	0,340	0,177	0,321	-0,119	0,375	0,329	0,549	0,387
shum2m _1_-1	0,336	0,276	-0,310	0,209	-0,465	-0,332	0,417	0,215
shum2m _0_-1	0,335	0,311	-0,299	-0,108	-0,045	-0,109	0,032	-0,465
shum2m _-1_-1	0,333	0,338	-0,300	-0,404	0,360	0,004	-0,492	0,190

Variable	PC9
shum2m _1_1	0,040
shum2m _0_1	-0,047
shum2m _-1_1	0,005
shum2m _1_0	-0,261
shum2m _0_0	0,440
shum2m _-1_0	-0,190
shum2m _1_-1	0,352
shum2m _0_-1	-0,678
shum2m _-1_-1	0,339

### Principal Component Analysis: temp2m \_1\_1; temp2m \_0\_1; temp2m \_-1\_1; temp2m \_1

#### Eigenanalysis of the Correlation Matrix

Eigenvalue	7,0625	1,8772	0,0425	0,0152	0,0026	0,0000	0,0000	0,0000
Proportion	0,785	0,209	0,005	0,002	0,000	0,000	0,000	0,000
Cumulative	0,785	0,993	0,998	1,000	1,000	1,000	1,000	1,000

Eigenvalue	0,0000
Proportion	0,000
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
temp2m _1_1	0,110	-0,690	0,700	-0,148	-0,020	0,001	0,000	0,000
temp2m _0_1	0,056	-0,714	-0,686	0,126	0,019	-0,001	-0,000	-0,000
temp2m _-1_1	0,373	0,056	-0,195	-0,897	-0,112	-0,057	0,000	-0,000
temp2m _1_0	0,375	0,042	0,011	0,213	-0,546	0,179	0,422	-0,537
temp2m _0_0	0,376	0,042	0,021	0,164	0,121	-0,389	-0,459	-0,125
temp2m _-1_0	0,376	0,042	0,020	0,113	0,474	-0,272	0,682	0,282
temp2m _1_-1	0,375	0,042	0,012	0,208	-0,480	0,122	-0,160	0,732
temp2m _0_-1	0,376	0,042	0,020	0,166	0,084	-0,357	-0,288	-0,272
temp2m _-1_-1	0,376	0,043	0,007	0,051	0,461	0,773	-0,196	-0,080

Variable	PC9
temp2m _1_1	0,000
temp2m _0_1	-0,000
temp2m _-1_1	0,000
temp2m _1_0	0,127
temp2m _0_0	0,662
temp2m _-1_0	0,035
temp2m _1_-1	-0,081
temp2m _0_-1	-0,733

temp2m \_-1\_-1 -0,010

Principal Component Analysis: uwind10m \_1\_; uwind10m \_0\_; uwind10m \_-1\_; uwind10

Eigenanalysis of the Correlation Matrix

Eigenvalue	8,9328	0,0372	0,0189	0,0059	0,0033	0,0012	0,0004	0,0002
Proportion	0,993	0,004	0,002	0,001	0,000	0,000	0,000	0,000
Cumulative	0,993	0,997	0,999	0,999	1,000	1,000	1,000	1,000
Eigenvalue	0,0000							
Proportion	0,000							
Cumulative	1,000							

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
uwind10m _1_1	0,333	-0,017	-0,760	-0,223	0,219	-0,384	-0,191	-0,174
uwind10m _0_1	0,333	-0,466	-0,233	-0,216	-0,591	0,388	0,135	0,224
uwind10m _-1_1	0,332	-0,643	0,286	0,434	0,115	-0,406	0,077	-0,148
uwind10m _1_0	0,334	0,238	-0,208	0,465	0,334	0,198	0,331	0,536
uwind10m _0_0	0,334	0,058	0,068	-0,077	0,296	0,504	0,249	-0,540
uwind10m _-1_0	0,334	-0,108	0,331	-0,326	0,377	0,226	-0,588	0,214
uwind10m _1_-1	0,333	0,361	0,036	0,433	-0,428	0,003	-0,533	-0,041
uwind10m _0_-1	0,334	0,329	0,160	-0,044	-0,241	-0,110	0,223	-0,403
uwind10m _-1_-1	0,334	0,243	0,319	-0,445	-0,081	-0,424	0,297	0,331

Variable	PC9
uwind10m _1_1	-0,016
uwind10m _0_1	0,044
uwind10m _-1_1	-0,028
uwind10m _1_0	0,159
uwind10m _0_0	-0,423
uwind10m _-1_0	0,276
uwind10m _1_-1	-0,317
uwind10m _0_-1	0,686
uwind10m _-1_-1	-0,381

Principal Component Analysis: vwind10m \_1\_; vwind10m \_0\_; vwind10m \_-1\_; vwind10

Eigenanalysis of the Correlation Matrix

Eigenvalue	8,7238	0,1566	0,0786	0,0221	0,0096	0,0061	0,0019	0,0010
Proportion	0,969	0,017	0,009	0,002	0,001	0,001	0,000	0,000
Cumulative	0,969	0,987	0,995	0,998	0,999	1,000	1,000	1,000

Eigenvalue	0,0001
Proportion	0,000
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
vwind10m _1_1	0,326	-0,629	-0,345	-0,289	0,084	-0,254	0,223	0,399
vwind10m _0_1	0,333	-0,401	0,254	-0,138	0,340	0,516	0,020	-0,490
vwind10m _-1_1	0,333	-0,216	0,519	0,028	-0,621	-0,026	-0,391	0,167
vwind10m _1_0	0,335	-0,100	-0,409	0,513	-0,065	-0,364	-0,261	-0,396
vwind10m _0_0	0,336	0,153	0,158	0,478	0,516	0,060	-0,155	0,324
vwind10m _-1_0	0,335	0,217	0,402	0,113	-0,020	-0,320	0,704	-0,015
vwind10m _1_-1	0,334	0,232	-0,403	0,070	-0,444	0,467	0,322	-0,109
vwind10m _0_-1	0,335	0,342	-0,175	-0,198	0,081	0,280	-0,219	0,445
vwind10m _-1_-1	0,333	0,382	-0,006	-0,591	0,127	-0,364	-0,238	-0,319

Variable	PC9
vwind10m _1_1	-0,124
vwind10m _0_1	0,153
vwind10m _-1_1	-0,072
vwind10m _1_0	0,292
vwind10m _0_0	-0,459
vwind10m _-1_0	0,260
vwind10m _1_-1	-0,368
vwind10m _0_-1	0,608

vwindl0m \_-1\_-1 -0,294

**Principal Component Analysis: sur\_vis\_1\_1; sur\_vis\_0\_1; sur\_vis\_-1\_; sur\_vis**

Eigenanalysis of the Correlation Matrix

Eigenvalue	4,3898	2,0894	0,9805	0,7040	0,3743	0,2498	0,1014	0,0774
Proportion	0,488	0,232	0,109	0,078	0,042	0,028	0,011	0,009
Cumulative	0,488	0,720	0,829	0,907	0,949	0,976	0,988	0,996

Eigenvalue	0,0334
Proportion	0,004
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
sur_vis_1_1	0,393	0,179	0,167	0,085	-0,701	0,371	0,101	0,117
sur_vis_0_1	0,334	0,409	-0,191	0,334	0,096	0,257	0,121	-0,450
sur_vis_-1_1	0,328	0,411	-0,189	0,224	0,386	-0,245	0,190	0,579
sur_vis_1_0	0,347	-0,200	-0,306	-0,587	-0,158	-0,237	0,526	-0,011
sur_vis_0_0	0,407	0,237	0,037	-0,273	-0,067	-0,447	-0,658	-0,242
sur_vis_-1_0	0,304	-0,018	0,614	-0,370	0,490	0,385	0,066	-0,012
sur_vis_1_-1	0,250	-0,391	-0,580	0,012	0,257	0,406	-0,229	-0,139
sur_vis_0_-1	0,325	-0,464	0,088	0,225	-0,104	0,029	-0,306	0,486
sur_vis_-1_-1	0,281	-0,409	0,289	0,473	0,068	-0,408	0,281	-0,368

Variable	PC9
sur_vis_1_1	0,356
sur_vis_0_1	-0,530
sur_vis_-1_1	0,237
sur_vis_1_0	-0,208
sur_vis_0_0	0,080
sur_vis_-1_0	0,023
sur_vis_1_-1	0,382
sur_vis_0_-1	-0,529
sur_vis_-1_-1	0,246

Lampiran 7. Output Reduksi PCA Stamet Soetta

Principal Component Analysis: hujan_total ; hujan_total ; hujan_total ; hujan_t								
Eigenanalysis of the Correlation Matrix								
Eigenvalue	6,7016	1,2401	0,4815	0,2608	0,1444	0,1071	0,0272	0,0232
Proportion	0,745	0,138	0,053	0,029	0,016	0,012	0,003	0,003
Cumulative	0,745	0,882	0,936	0,965	0,981	0,993	0,996	0,998
Eigenvalue	0,0143							
Proportion	0,002							
Cumulative	1,000							
Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	
hujan_total_1_1	0,317	-0,370	-0,495	0,162	-0,077	-0,512	0,136	
hujan_total_0_1	0,335	-0,422	-0,166	0,105	-0,118	0,014	-0,224	
hujan_total_-1_1	0,348	-0,292	0,095	-0,045	-0,456	0,632	0,175	
hujan_total_1_0	0,335	-0,218	0,137	-0,747	0,494	0,006	0,042	
hujan_total_0_0	0,355	-0,001	0,461	0,299	0,244	-0,315	0,009	
hujan_total_-1_0	0,356	0,080	0,411	0,420	0,141	0,129	-0,099	
hujan_total_1_-1	0,322	0,421	0,072	-0,328	-0,512	-0,274	-0,513	
hujan_total_0_-1	0,323	0,470	-0,099	-0,059	-0,125	-0,078	0,740	
hujan_total_-1_-1	0,304	0,384	-0,552	0,159	0,413	0,377	-0,280	
Variable	PC8	PC9						
hujan_total_1_1	-0,337	-0,303						
hujan_total_0_1	0,369	0,686						
hujan_total_-1_1	0,043	-0,377						
hujan_total_1_0	-0,131	0,020						
hujan_total_0_0	0,541	-0,348						
hujan_total_-1_0	-0,644	0,247						
hujan_total_1_-1	-0,047	-0,056						
hujan_total_0_-1	0,090	0,290						
hujan_total_-1_-1	0,116	-0,161						
Principal Component Analysis: dew2m_1_1; dew2m_0_1; dew2m_-1_1; dew2m_1_0;								
Eigenanalysis of the Correlation Matrix								
Eigenvalue	8,4027	0,5104	0,0514	0,0190	0,0077	0,0061	0,0015	0,0006
Proportion	0,934	0,057	0,006	0,002	0,001	0,001	0,000	0,000
Cumulative	0,934	0,990	0,996	0,998	0,999	1,000	1,000	1,000
Eigenvalue	0,0006							
Proportion	0,000							
Cumulative	1,000							
Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
dew2m_1_1	0,324	-0,462	0,313	-0,291	0,388	0,396	-0,230	-0,296
dew2m_0_1	0,326	-0,458	-0,070	0,015	-0,053	0,189	0,331	0,626
dew2m_-1_1	0,327	-0,411	-0,482	0,138	-0,517	-0,186	-0,168	-0,362
dew2m_1_0	0,342	-0,083	0,469	-0,048	0,061	-0,793	0,091	0,066
dew2m_0_0	0,340	0,185	0,114	0,615	0,172	0,172	0,459	-0,409
dew2m_-1_0	0,339	0,202	-0,375	0,328	0,473	-0,085	-0,467	0,339
dew2m_1_-1	0,334	0,330	0,388	0,021	-0,433	0,207	-0,499	0,046
dew2m_0_-1	0,335	0,335	0,006	-0,194	-0,306	0,232	0,306	0,217
dew2m_-1_-1	0,333	0,324	-0,375	-0,607	0,202	-0,103	0,170	-0,229
Variable	PC9							
dew2m_1_1	0,224							
dew2m_0_1	-0,372							
dew2m_-1_1	0,108							
dew2m_1_0	0,091							
dew2m_0_0	-0,143							
dew2m_-1_0	0,177							
dew2m_1_-1	-0,384							
dew2m_0_-1	0,671							
dew2m_-1_-1	-0,376							



**Principal Component Analysis: wind\_gust\_1; wind\_gust\_0; wind\_gust\_-1; wind\_gust**

Eigenanalysis of the Correlation Matrix

Eigenvalue	8,0366	0,7107	0,1450	0,0506	0,0263	0,0169	0,0062	0,0056
Proportion	0,893	0,079	0,016	0,006	0,003	0,002	0,001	0,001
Cumulative	0,893	0,972	0,988	0,994	0,997	0,998	0,999	1,000

Eigenvalue	0,0020
Proportion	0,000
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
wind_gust_1_1	0,314	-0,514	0,180	-0,477	0,292	-0,324	-0,213	-0,121
wind_gust_0_1	0,334	-0,363	-0,196	-0,210	-0,148	-0,046	0,309	-0,018
wind_gust_-1_1	0,340	-0,163	-0,560	0,039	-0,544	0,231	-0,074	0,050
wind_gust_1_0	0,341	-0,237	0,302	0,279	0,265	0,699	-0,247	0,174
wind_gust_0_0	0,350	-0,028	0,037	0,457	0,253	-0,251	0,660	0,179
wind_gust_-1_0	0,340	0,234	-0,385	0,364	0,305	-0,340	-0,509	-0,194
wind_gust_1_-1	0,340	0,177	0,526	0,142	-0,457	-0,068	0,008	-0,586
wind_gust_0_-1	0,333	0,356	0,263	-0,204	-0,246	-0,232	-0,191	0,704
wind_gust_-1_-1	0,308	0,554	-0,170	-0,495	0,317	0,338	0,246	-0,209

Variable	PC9
wind_gust_1_1	-0,355
wind_gust_0_1	0,744
wind_gust_-1_1	-0,431
wind_gust_1_0	0,093
wind_gust_0_0	-0,267
wind_gust_-1_0	0,209
wind_gust_1_-1	0,000
wind_gust_0_-1	0,069
wind_gust_-1_-1	-0,075

**Principal Component Analysis: hi\_cloud\_1; hi\_cloud\_0; hi\_cloud\_-1; hi\_cloud**

Eigenanalysis of the Correlation Matrix

Eigenvalue	8,1867	0,4565	0,2165	0,0510	0,0431	0,0246	0,0096	0,0081
Proportion	0,910	0,051	0,024	0,006	0,005	0,003	0,001	0,001
Cumulative	0,910	0,960	0,984	0,990	0,995	0,998	0,999	1,000

Eigenvalue	0,0039
Proportion	0,000
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
hi_cloud_1_1	0,329	-0,420	-0,239	-0,561	0,071	0,211	-0,046	-0,447
hi_cloud_0_1	0,333	-0,414	0,055	-0,101	0,119	-0,565	-0,381	0,253
hi_cloud_-1_1	0,326	-0,314	0,574	0,335	0,425	0,188	0,320	0,053
hi_cloud_1_0	0,341	-0,144	-0,373	0,058	-0,218	0,491	0,305	0,381
hi_cloud_0_0	0,346	-0,017	-0,119	0,162	-0,517	-0,417	0,276	0,186
hi_cloud_-1_0	0,338	0,135	0,417	0,165	-0,540	0,197	-0,339	-0,414
hi_cloud_1_-1	0,335	0,237	-0,426	0,463	0,340	0,133	-0,494	0,002
hi_cloud_0_-1	0,333	0,417	-0,144	-0,014	0,257	-0,340	0,459	-0,441
hi_cloud_-1_-1	0,319	0,533	0,284	-0,540	0,109	0,114	-0,106	0,433

Variable	PC9
hi_cloud_1_1	0,303
hi_cloud_0_1	-0,401
hi_cloud_-1_1	0,179
hi_cloud_1_0	-0,440
hi_cloud_0_0	0,537
hi_cloud_-1_0	-0,220
hi_cloud_1_-1	0,242
hi_cloud_0_-1	-0,329
hi_cloud_-1_-1	0,135

**Principal Component Analysis: low\_cloud\_1; low\_cloud\_0; low\_cloud\_-; low\_clo**

Eigenanalysis of the Correlation Matrix

Eigenvalue	6,1584	1,3426	0,6050	0,4091	0,2265	0,1117	0,0662	0,0554
Proportion	0,684	0,149	0,067	0,045	0,025	0,012	0,007	0,006
Cumulative	0,684	0,833	0,901	0,946	0,971	0,984	0,991	0,997

Eigenvalue	0,0251
Proportion	0,003
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
low_cloud_1_1	0,293	-0,479	0,277	-0,408	-0,353	-0,266	-0,258	-0,363
low_cloud_0_1	0,345	-0,383	-0,036	-0,241	0,207	-0,181	0,560	0,290
low_cloud_-1_1	0,334	-0,296	-0,385	0,007	0,617	0,332	-0,261	-0,136
low_cloud_1_0	0,354	-0,096	0,468	0,239	-0,195	0,599	-0,233	0,255
low_cloud_0_0	0,380	-0,014	-0,085	0,404	-0,255	-0,177	0,268	0,366
low_cloud_-1_0	0,348	0,072	-0,492	0,384	-0,336	-0,117	-0,073	-0,413
low_cloud_1_-1	0,313	0,372	0,498	0,163	0,364	-0,102	0,305	-0,499
low_cloud_0_-1	0,340	0,404	0,026	-0,161	0,195	-0,459	-0,530	0,382
low_cloud_-1_-1	0,282	0,470	-0,237	-0,596	-0,255	0,401	0,215	-0,026

Variable	PC9
low_cloud_1_1	-0,218
low_cloud_0_1	0,448
low_cloud_-1_1	-0,273
low_cloud_1_0	0,270
low_cloud_0_0	-0,619
low_cloud_-1_0	0,426
low_cloud_1_-1	-0,063
low_cloud_0_-1	0,138
low_cloud_-1_-1	-0,126

**Principal Component Analysis: mid\_cloud\_1; mid\_cloud\_0; mid\_cloud\_-; mid\_clo**

Eigenanalysis of the Correlation Matrix

Eigenvalue	8,1583	0,5440	0,1554	0,0541	0,0405	0,0241	0,0109	0,0082
Proportion	0,906	0,060	0,017	0,006	0,005	0,003	0,001	0,001
Cumulative	0,906	0,967	0,984	0,990	0,995	0,997	0,999	1,000

Eigenvalue	0,0045
Proportion	0,000
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
mid_cloud_1_1	0,328	-0,430	0,200	-0,310	0,426	0,205	-0,476	-0,129
mid_cloud_0_1	0,335	-0,375	-0,064	-0,208	-0,050	0,329	0,399	-0,189
mid_cloud_-1_1	0,334	-0,282	-0,442	-0,291	-0,463	-0,277	0,069	0,368
mid_cloud_1_0	0,338	-0,209	0,381	0,395	0,274	-0,481	0,115	0,428
mid_cloud_0_0	0,346	-0,010	-0,059	0,589	-0,052	0,408	0,332	-0,152
mid_cloud_-1_0	0,341	0,113	-0,453	0,349	-0,037	-0,234	-0,517	-0,367
mid_cloud_1_-1	0,333	0,250	0,568	-0,203	-0,447	-0,258	0,040	-0,431
mid_cloud_0_-1	0,328	0,453	0,131	-0,052	-0,166	0,475	-0,309	0,532
mid_cloud_-1_-1	0,318	0,520	-0,259	-0,326	0,543	-0,168	0,347	-0,050

Variable	PC9
mid_cloud_1_1	-0,323
mid_cloud_0_1	0,628
mid_cloud_-1_1	-0,314
mid_cloud_1_0	0,194
mid_cloud_0_0	-0,476
mid_cloud_-1_0	0,293
mid_cloud_1_-1	-0,092
mid_cloud_0_-1	0,187
mid_cloud_-1_-1	-0,098

**Principal Component Analysis: tot\_cloud\_1; tot\_cloud\_0; tot\_cloud\_-; tot\_clo**

Eigenanalysis of the Correlation Matrix

Eigenvalue	8,1743	0,5043	0,1878	0,0489	0,0401	0,0217	0,0111	0,0083
Proportion	0,908	0,056	0,021	0,005	0,004	0,002	0,001	0,001
Cumulative	0,908	0,964	0,985	0,991	0,995	0,997	0,999	1,000

Eigenvalue	0,0035
Proportion	0,000
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
tot_cloud_1_1	0,329	-0,421	-0,214	0,478	0,238	0,018	-0,443	-0,244
tot_cloud_0_1	0,335	-0,382	0,057	0,200	-0,114	0,403	0,362	-0,168
tot_cloud_-1_1	0,332	-0,276	0,478	-0,006	-0,589	-0,185	-0,013	0,349
tot_cloud_1_0	0,341	-0,183	-0,336	-0,146	0,327	-0,544	0,101	0,479
tot_cloud_0_0	0,346	-0,033	-0,015	-0,454	0,332	0,432	0,365	0,059
tot_cloud_-1_0	0,339	0,101	0,466	-0,413	0,215	-0,222	-0,347	-0,462
tot_cloud_1_-1	0,332	0,246	-0,543	-0,145	-0,511	-0,203	0,149	-0,423
tot_cloud_0_-1	0,331	0,419	-0,154	-0,009	-0,121	0,452	-0,518	0,405
tot_cloud_-1_-1	0,314	0,565	0,270	0,559	0,211	-0,162	0,344	-0,000

Variable	PC9
tot_cloud_1_1	-0,357
tot_cloud_0_1	0,603
tot_cloud_-1_1	-0,286
tot_cloud_1_0	0,273
tot_cloud_0_0	-0,489
tot_cloud_-1_0	0,238
tot_cloud_1_-1	-0,098
tot_cloud_0_-1	0,199
tot_cloud_-1_-1	-0,086

**Principal Component Analysis: mslp\_1\_1; mslp\_0\_1; mslp\_-1\_1; mslp\_1\_0; mslp**

Eigenanalysis of the Correlation Matrix

Eigenvalue	8,9363	0,0365	0,0207	0,0030	0,0016	0,0011	0,0005	0,0002
Proportion	0,993	0,004	0,002	0,000	0,000	0,000	0,000	0,000
Cumulative	0,993	0,997	0,999	1,000	1,000	1,000	1,000	1,000

Eigenvalue	0,0001
Proportion	0,000
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
mslp_1_1	0,332	-0,643	-0,055	0,129	0,130	0,405	0,120	-0,244
mslp_0_1	0,333	-0,492	0,072	0,227	-0,049	-0,122	-0,152	0,311
mslp_-1_1	0,334	-0,032	0,400	0,047	-0,527	-0,553	0,161	-0,022
mslp_1_0	0,334	-0,010	0,136	-0,724	0,307	-0,113	0,303	-0,282
mslp_0_0	0,334	0,134	0,202	-0,270	0,245	0,084	-0,607	0,469
mslp_-1_0	0,333	0,395	0,323	0,146	-0,327	0,564	-0,122	-0,329
mslp_1_-1	0,333	0,144	-0,639	-0,215	-0,406	0,181	0,268	0,379
mslp_0_-1	0,333	0,143	-0,505	0,181	0,113	-0,374	-0,426	-0,497
mslp_-1_-1	0,334	0,358	0,066	0,482	0,514	-0,070	0,456	0,213

Variable	PC9
mslp_1_1	-0,449
mslp_0_1	0,673
mslp_-1_1	-0,340
mslp_1_0	0,258
mslp_0_0	-0,318
mslp_-1_0	0,243
mslp_1_-1	-0,030
mslp_0_-1	0,007
mslp_-1_-1	-0,044

**Principal Component Analysis: sur\_press\_1; sur\_press\_0; sur\_press\_-; sur\_pre**

Eigenanalysis of the Correlation Matrix

Eigenvalue	8,9668	0,0238	0,0076	0,0009	0,0005	0,0002	0,0001	0,0001
Proportion	0,996	0,003	0,001	0,000	0,000	0,000	0,000	0,000
Cumulative	0,996	0,999	1,000	1,000	1,000	1,000	1,000	1,000

Eigenvalue	0,0000
Proportion	0,000
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
sur_press_1_1	0,333	-0,400	-0,302	0,576	0,049	-0,201	-0,379	-0,270
sur_press_0_1	0,333	-0,390	0,023	0,080	-0,197	-0,224	0,197	0,654
sur_press_-1_1	0,333	-0,386	0,387	-0,291	-0,494	0,181	0,190	-0,364
sur_press_1_0	0,334	-0,074	-0,390	-0,016	0,348	0,548	0,395	-0,229
sur_press_0_0	0,334	-0,027	-0,034	-0,341	0,510	-0,238	0,177	0,257
sur_press_-1_0	0,334	0,017	0,404	-0,273	0,338	0,080	-0,627	-0,057
sur_press_1_-1	0,333	0,377	-0,479	-0,234	-0,458	0,229	-0,346	0,241
sur_press_0_-1	0,333	0,426	-0,065	-0,069	-0,100	-0,630	0,239	-0,399
sur_press_-1_-1	0,333	0,458	0,456	0,571	0,002	0,255	0,155	0,167

Variable	PC9
sur_press_1_1	-0,215
sur_press_0_1	0,418
sur_press_-1_1	-0,247
sur_press_1_0	0,318
sur_press_0_0	-0,597
sur_press_-1_0	0,365
sur_press_1_-1	-0,150
sur_press_0_-1	0,275
sur_press_-1_-1	-0,169

**Principal Component Analysis: rhum2m\_1\_1; rhum2m\_0\_1; rhum2m\_-1\_1; rhum2m\_1**

Eigenanalysis of the Correlation Matrix

Eigenvalue	7,6884	1,1411	0,0804	0,0469	0,0237	0,0130	0,0032	0,0020
Proportion	0,854	0,127	0,009	0,005	0,003	0,001	0,000	0,000
Cumulative	0,854	0,981	0,990	0,995	0,998	0,999	1,000	1,000

Eigenvalue	0,0013
Proportion	0,000
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
rhum2m_1_1	0,311	-0,461	0,242	-0,306	0,420	-0,411	0,060	0,128
rhum2m_0_1	0,314	-0,460	-0,058	0,005	-0,037	-0,178	-0,003	-0,241
rhum2m_-1_1	0,314	-0,435	-0,465	0,216	-0,512	0,179	0,004	0,055
rhum2m_1_0	0,353	-0,101	0,494	-0,073	0,062	0,763	-0,091	0,135
rhum2m_0_0	0,348	0,206	0,113	0,593	0,179	-0,192	-0,591	-0,200
rhum2m_-1_0	0,345	0,229	-0,428	0,257	0,522	0,142	0,480	0,219
rhum2m_1_-1	0,340	0,287	0,381	0,080	-0,358	-0,184	0,563	-0,396
rhum2m_0_-1	0,338	0,314	0,044	-0,181	-0,347	-0,288	-0,150	0,700
rhum2m_-1_-1	0,334	0,316	-0,366	-0,630	0,049	0,108	-0,263	-0,411

Variable	PC9
rhum2m_1_1	-0,416
rhum2m_0_1	0,772
rhum2m_-1_1	-0,389
rhum2m_1_0	0,054
rhum2m_0_0	-0,115
rhum2m_-1_0	0,090
rhum2m_1_-1	-0,121
rhum2m_0_-1	0,190
rhum2m_-1_-1	-0,072

**Principal Component Analysis: shum2m \_1\_1; shum2m \_0\_1; shum2m \_-1\_1; shum2m \_1**

Eigenanalysis of the Correlation Matrix

Eigenvalue	8,3870	0,5251	0,0509	0,0198	0,0081	0,0062	0,0016	0,0007
Proportion	0,932	0,058	0,006	0,002	0,001	0,001	0,000	0,000
Cumulative	0,932	0,990	0,996	0,998	0,999	1,000	1,000	1,000

Eigenvalue	0,0006
Proportion	0,000
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
shum2m _1_1	0,324	-0,461	0,311	-0,276	0,416	0,384	-0,206	0,186
shum2m _0_1	0,325	-0,459	-0,068	0,029	-0,043	0,193	0,288	-0,421
shum2m _-1_1	0,327	-0,413	-0,472	0,147	-0,529	-0,159	-0,156	0,286
shum2m _1_0	0,342	-0,082	0,474	-0,090	0,005	-0,787	0,097	-0,106
shum2m _0_0	0,340	0,191	0,108	0,613	0,190	0,129	0,484	0,416
shum2m _-1_0	0,340	0,201	-0,381	0,311	0,480	-0,140	-0,466	-0,366
shum2m _1_-1	0,333	0,333	0,391	0,049	-0,401	0,230	-0,520	0,135
shum2m _0_-1	0,335	0,334	0,003	-0,182	-0,298	0,271	0,303	-0,488
shum2m _-1_-1	0,333	0,317	-0,376	-0,620	0,171	-0,089	0,168	0,362

Variable	PC9
shum2m _1_1	0,335
shum2m _0_1	-0,616
shum2m _-1_1	0,258
shum2m _1_0	0,056
shum2m _0_0	0,023
shum2m _-1_0	0,041
shum2m _1_-1	-0,347
shum2m _0_-1	0,501
shum2m _-1_-1	-0,256

**Principal Component Analysis: temp2m \_1\_1; temp2m \_0\_1; temp2m \_-1\_1; temp2m \_1**

Eigenanalysis of the Correlation Matrix

Eigenvalue	5,2069	3,5476	0,1358	0,0846	0,0109	0,0069	0,0066	0,0007
Proportion	0,579	0,394	0,015	0,009	0,001	0,001	0,001	0,000
Cumulative	0,579	0,973	0,988	0,997	0,998	0,999	1,000	1,000

Eigenvalue	0,0000
Proportion	0,000
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
temp2m _1_1	-0,268	-0,419	0,020	-0,049	0,062	-0,262	0,445	0,648
temp2m _0_1	-0,266	-0,422	0,029	-0,027	-0,048	0,149	-0,271	0,162
temp2m _-1_1	-0,264	-0,423	0,033	-0,018	-0,093	0,316	-0,567	-0,070
temp2m _1_0	-0,267	-0,420	0,011	-0,050	0,065	-0,212	0,389	-0,741
temp2m _0_0	0,366	-0,270	0,290	0,565	0,615	0,107	-0,012	0,004
temp2m _-1_0	0,353	-0,287	-0,498	0,489	-0,517	-0,196	-0,019	-0,004
temp2m _1_-1	0,395	-0,207	0,461	-0,196	-0,441	0,485	0,347	0,000
temp2m _0_-1	0,399	-0,198	0,333	-0,374	-0,033	-0,647	-0,362	-0,010
temp2m _-1_-1	0,377	-0,232	-0,585	-0,508	0,375	0,246	0,054	0,012

Variable	PC9
temp2m _1_1	0,243
temp2m _0_1	-0,791
temp2m _-1_1	0,561
temp2m _1_0	-0,013
temp2m _0_0	-0,000
temp2m _-1_0	-0,000
temp2m _1_-1	0,000
temp2m _0_-1	-0,001
temp2m _-1_-1	0,001

**Principal Component Analysis: uwind10m \_1\_; uwind10m \_0\_; uwind10m \_-1\_; uwind10**

Eigenanalysis of the Correlation Matrix

Eigenvalue	8,8004	0,1321	0,0482	0,0116	0,0035	0,0020	0,0016	0,0004
Proportion	0,978	0,015	0,005	0,001	0,000	0,000	0,000	0,000
Cumulative	0,978	0,992	0,998	0,999	1,000	1,000	1,000	1,000

Eigenvalue	0,0002
Proportion	0,000
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
uwind10m _1_1	0,332	-0,441	-0,162	-0,556	-0,232	0,336	-0,243	0,264
uwind10m _0_1	0,333	-0,405	0,078	-0,040	0,071	0,147	0,139	-0,531
uwind10m _-1_1	0,333	-0,334	0,356	0,388	0,550	0,095	0,126	0,178
uwind10m _1_0	0,335	-0,141	-0,350	-0,123	0,117	-0,797	0,160	0,219
uwind10m _0_0	0,337	0,018	-0,076	0,319	-0,596	-0,067	0,269	-0,387
uwind10m _-1_0	0,335	0,102	0,355	0,345	-0,388	-0,072	-0,418	0,409
uwind10m _1_-1	0,332	0,307	-0,590	0,251	0,305	0,205	-0,466	-0,180
uwind10m _0_-1	0,332	0,451	-0,089	-0,118	0,025	0,360	0,617	0,344
uwind10m _-1_-1	0,330	0,449	0,481	-0,476	0,158	-0,200	-0,188	-0,318

Variable	PC9
uwind10m _1_1	-0,253
uwind10m _0_1	0,624
uwind10m _-1_1	-0,376
uwind10m _1_0	0,091
uwind10m _0_0	-0,444
uwind10m _-1_0	0,367
uwind10m _1_-1	-0,021
uwind10m _0_-1	0,186
uwind10m _-1_-1	-0,175

**Principal Component Analysis: vwind10m \_1\_; vwind10m \_0\_; vwind10m \_-1\_; vwind10**

Eigenanalysis of the Correlation Matrix

Eigenvalue	7,8208	0,7793	0,2855	0,0625	0,0253	0,0139	0,0074	0,0040
Proportion	0,869	0,087	0,032	0,007	0,003	0,002	0,001	0,000
Cumulative	0,869	0,956	0,987	0,994	0,997	0,999	0,999	1,000

Eigenvalue	0,0012
Proportion	0,000
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
vwind10m _1_1	0,322	0,463	0,031	-0,534	-0,159	0,316	-0,281	-0,279
vwind10m _0_1	0,338	0,335	0,240	-0,138	0,160	0,184	0,179	0,325
vwind10m _-1_1	0,339	0,163	0,498	0,168	0,439	-0,379	0,225	0,020
vwind10m _1_0	0,338	0,258	-0,389	0,059	-0,473	-0,571	0,294	-0,132
vwind10m _0_0	0,353	0,023	-0,124	0,500	-0,294	0,413	-0,027	0,507
vwind10m _-1_0	0,346	-0,188	0,280	0,440	-0,120	0,031	-0,427	-0,540
vwind10m _1_-1	0,336	-0,093	-0,591	-0,038	0,571	-0,148	-0,407	0,116
vwind10m _0_-1	0,328	-0,430	-0,183	-0,120	0,147	0,378	0,611	-0,348
vwind10m _-1_-1	0,297	-0,592	0,258	-0,452	-0,289	-0,246	-0,174	0,340

Variable	PC9
vwind10m _1_1	-0,336
vwind10m _0_1	0,707
vwind10m _-1_1	-0,442
vwind10m _1_0	0,102
vwind10m _0_0	-0,307
vwind10m _-1_0	0,290
vwind10m _1_-1	0,036
vwind10m _0_-1	-0,040
vwind10m _-1_-1	-0,020

**Principal Component Analysis: sur\_vis\_1\_1; sur\_vis\_0\_1; sur\_vis\_1\_0; sur\_vis\_0\_0**

Eigenanalysis of the Correlation Matrix

Eigenvalue	6,3185	1,6970	0,4209	0,2152	0,1377	0,1029	0,0529	0,0380
Proportion	0,702	0,189	0,047	0,024	0,015	0,011	0,006	0,004
Cumulative	0,702	0,891	0,937	0,961	0,977	0,988	0,994	0,998

Eigenvalue	0,0170
Proportion	0,002
Cumulative	1,000

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
sur_vis_1_1	0,327	0,337	-0,075	-0,679	0,233	0,460	-0,055	0,155
sur_vis_0_1	0,354	0,319	-0,058	-0,049	-0,357	-0,024	0,277	-0,411
sur_vis_1_0	0,350	0,302	-0,032	0,095	-0,622	-0,270	-0,265	0,282
sur_vis_0_0	0,361	0,222	0,034	0,092	0,577	-0,611	0,134	0,269
sur_vis_1_1	0,388	-0,009	0,098	0,291	0,195	0,137	0,203	-0,598
sur_vis_1_0	0,375	-0,080	-0,064	0,538	0,141	0,469	-0,420	0,231
sur_vis_1_1	0,271	-0,393	0,775	-0,271	-0,074	-0,111	-0,248	-0,065
sur_vis_0_1	0,297	-0,486	-0,096	-0,004	-0,185	0,140	0,659	0,420
sur_vis_1_1	0,250	-0,497	-0,604	-0,267	0,023	-0,267	-0,344	-0,252

Variable	PC9
sur_vis_1_1	0,140
sur_vis_0_1	-0,628
sur_vis_1_0	0,408
sur_vis_0_0	-0,115
sur_vis_1_1	0,547
sur_vis_1_0	-0,300
sur_vis_1_1	-0,119
sur_vis_0_1	0,046
sur_vis_1_1	-0,024

Lampiran 8. Output CART Stamet Juanda

=====

TREE SEQUENCE

=====

Dependent variable: CURAHHUJ

Terminal Tree	Nodes	Cross-Validated Relative Cost	Resubstitution Relative Cost	Complexity Parameter
1	45	0.887 +/- 0.045	0.118	0.000
15	15	0.830 +/- 0.046	0.291	0.009
16	13	0.820 +/- 0.046	0.318	0.010
17	11	0.827 +/- 0.046	0.350	0.012
18	9	0.813 +/- 0.054	0.388	0.014
19	8	0.774 +/- 0.061	0.410	0.017
20**	6	0.730 +/- 0.060	0.485	0.028
21	4	0.734 +/- 0.057	0.568	0.031
22	3	0.757 +/- 0.061	0.625	0.043
23	2	0.840 +/- 0.046	0.796	0.129
24	1	1.000 +/- 0.000	1.000	0.153

Initial misclassification cost = 0.750

Initial class assignment = 1

=====

VARIABLE IMPORTANCE

=====

	Relative Importance	Number Of Categories	Minimum Category
HUJAN_1T	100.000		
RHUM2M1	90.157		
SHUM2M	70.608		
V1WIND10	63.622		
DEW2M	57.116		
SUR_PRES	47.054		
TOT_CLOU	45.793		
TEMP2M2	38.137		
HI_CLOUD	31.918		
MID_CLOU	31.063		
LOW_1CLO	30.830		
UWIND10M	22.425		
HUJAN_2T	21.050		
LOW_2CLO	16.144		
V2WIND10	12.535		
SUR_VIS1	10.670		
SUR_VIS2	9.855		
WIND_GUS	9.673		
TEMP2M1	0.000		
MSLP	0.000		





Lampiran 9. Output CART Stamet Ngurai Rai

=====				
TREE SEQUENCE				
=====				
Dependent variable: CURAHHUJ				
Terminal		Cross-Validated	Resubstitution	Complexity
Tree Nodes		Relative Cost	Relative Cost	Parameter
-----				
1	47	0.929 +/- 0.026	0.047	0.000
4**	33	0.919 +/- 0.026	0.070	0.002
18	11	0.983 +/- 0.026	0.249	0.016
19	9	0.993 +/- 0.025	0.295	0.018
20	8	1.022 +/- 0.012	0.320	0.020
21	7	1.034 +/- 0.011	0.350	0.024
22	6	1.040 +/- 0.009	0.383	0.026
23	5	0.945 +/- 0.054	0.428	0.036
24	4	0.998 +/- 0.057	0.514	0.069
25	3	1.021 +/- 0.054	0.628	0.091
26	2	0.962 +/- 0.022	0.750	0.098
27	1	1.000 +/- 0.000	1.000	0.200
Initial misclassification cost = 0.800				
Initial class assignment = 1				
=====				
VARIABLE IMPORTANCE				
=====				
	Relative	Number Of	Minimum	
	Importance	Categories	Category	
-----				
TEMP2M1	100.000			
SUR_PRES	81.632			
RHUM2M1	75.874			
MSLP	69.380			
LOW_CLOU	51.095			
HUJAN_TO	50.682			
SUR_VIS2	50.075			
HI_CLOUD	49.706			
VWIND10M	41.000			
UWIND10M	38.281			
TOT_CLOU	37.225			
SUR_VIS1	36.533			
RHUM2M2	32.952			
WIND_GUS	31.411			
TEMP2M2	26.903			
SUR_VIS3	25.027			
DEW2M	10.719			
SHUM2M	10.334			
MID_CLOU	4.150			





## Lampiran 10. Output CART Stamet Soetta

```
=====
TREE SEQUENCE
=====
```

Dependent variable: CURAHHUJ

Terminal Tree	Nodes	Cross-Validated Relative Cost	Resubstitution Relative Cost	Complexity Parameter
1	45	0.921 +/- 0.050	0.076	0.000
17	14	0.894 +/- 0.055	0.212	0.013
18	12	0.895 +/- 0.055	0.245	0.013
19	10	0.878 +/- 0.056	0.279	0.014
20	8	0.863 +/- 0.059	0.326	0.019
21	7	0.854 +/- 0.059	0.350	0.020
22**	6	0.850 +/- 0.059	0.375	0.020
23	4	0.918 +/- 0.057	0.460	0.034
24	3	0.924 +/- 0.053	0.583	0.099
25	2	0.999 +/- 0.012	0.750	0.133
26	1	1.000 +/- 0.000	1.000	0.200

Initial misclassification cost = 0.800

Initial class assignment = 1

```
=====
VARIABLE IMPORTANCE
=====
```

	Relative Importance	Number Of Categories	Minimum Category
MID_CLOU	100.000		
RHUM2M2	98.533		
TOT_CLOU	80.387		
TEMP2M1	44.247		
RHUM2M1	43.913		
WIND_GUS	43.674		
TEMP2M2	37.329		
SUR_VIS1	31.990		
HI_CLOUD	29.232		
V8_A	28.890		
VWIND10M	26.288		
HUJAN_2T	23.031		
UWIND10M	16.829		
LOW_CLOU	12.686		
HUJAN_1T	9.723		
SHUM2M	7.549		
SUR_VIS2	3.847		
SUR_PRES	1.464		
DEW2M	0.000		
MSLP	0.000		



### Lampiran 11. Program R Untuk Random Forests

```
#Install package : randomForest
#Random forests prediction of 'Curah Hujan'
library(randomForest) #load library
juandalearn=read.csv("D:/juandalearn.csv", sep=";", header=TRUE)
juandatest=read.csv("D:/juandatest.csv", sep=";", header=TRUE)

#grow tree for RF
fit.rf=randomForest(curahhujan~., data=juandalearn, mtry=3, ntree=5)

#accuracy prediction and calculation of learning dataset
fitrf.pred1=predict(fit.rf,juandalearn[-347,])
write.csv(fitrf.pred1, "D:\\rf\\rfjuandalearn.csv")
crosstabrf1=table(predicted=fitrf.pred1,observed=juandalearn[-347,"curahhujan"])
APERrf1=1-(sum(diag(crosstabrf1))/sum(crosstabrf1))
accuracyrf1=(sum(diag(crosstabrf1))/sum(crosstabrf1))
print(crosstabrf1)
print(APERrf1)
print(accuracyrf1)

#accuracy prediction and calculate of testing dataset
fitrf.pred2=predict(fit.rf,juandatest[-8,])
write.csv(fitrf.pred2,"D:\\rf\\rfjuandatest.csv")
crosstabrf2=table(predicted=fitrf.pred2,observed=juandatest[-8,"curahhujan"])
APERrf2= 1-(sum(diag(crosstabrf2))/sum(crosstabrf2))
accuracyrf2=(sum(diag(crosstabrf2))/sum(crosstabrf2))
print(crosstabrf2)
print(APERrf2)
print(accuracyrf2)
```

kombinasi jumlah pohon sebesar 10, 15, 20 dan 25 diperoleh dengan cara mengubah input (ntree=5) sesuai dengan jumlah pohon yang diinginkan





## SURAT PERNYATAAN

Saya yang bertanda tangan di bawah ini, mahasiswa Departemen Statistika FMIPA ITS

Nama : Nurhayati

NRP : 1315 105 023

menyatakan bahwa data yang digunakan dalam Tugas Akhir/Thesis ini merupakan data sekunder yang diambil dari penelitian/~~buku~~/Tugas Akhir/Thesis/publikasi lainnya yaitu:

Sumber : Data Penelitian Dr. Sutikno, M.Si dengan Judul  
"Pengembangan Model Prakiraan Cuaca melalui Model Output Statistik dengan Pendekatan Linear dan Non-Linear dalam Antisipasi Bencana Hidrometeorologi"

Keterangan : Data curah hujan harian dan *Output* data NWP WRF-EMS pada Stasiun Meteorologi Juanda, Ngurai Rai dan Soekarno-Hatta periode 2 Januari 2015 – 10 Desember 2016

Surat pernyataan ini dibuat dengan sebenarnya. Apabila terdapat pemalsuan data maka saya siap menerima sanksi sesuai aturan yang berlaku.

Mengetahui,  
Pembimbing Tugas Akhir



(Dr. Sutikno, M. Si)  
NIP. 19710313 199702 1 001

\*(coret yang tidak perlu)

Surabaya, 13 Juni 2017  
Mahasiswa



(Nurhayati)  
NRP. 1315 105 023



## **BIODATA PENULIS**



Penulis terlahir di Bangkalan, pada tanggal 17 Desember 1993 yang bernama lengkap Nurhayati. Penulis merupakan anak tunggal. Penulis menyelesaikan pendidikan formal di TK NU Muslimat Arosbaya, SDN 1 Arosbaya, SMPN 1 Arosbaya, SMAN 1 Bangkalan dan Diploma III Statistika ITS. Selanjutnya pada tahun 2015 penulis di terima di ITS tepatnya pada program studi Lintas Jalur Statistika dan terdaftar dengan

NRP 1315 105 023. Selama menjadi mahasiswa, penulis aktif mengikuti kegiatan kemahasiswaan baik kepanitiaan maupun organisasi. Organisasi kemahasiswaan yang pernah penulis ikuti yaitu sebagai staff ED WE&T ITS (2013/2014), staff Danus FORSIS ITS (2013/2014). Beberapa kegiatan kepanitiaan diantaranya POMITS, STATION, dan masih banyak lainnya. Apabila pembaca ingin berdiskusi mengenai Tugas Akhir ini atau semua yang berhubungan dengan penulis dapat mengirimkan email ke [nur\\_hyatie17@yahoo.com](mailto:nur_hyatie17@yahoo.com)

